MAR 7 1925

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. J. ASHBAUGH — W. W. CHARTERS — GEORGE D. STRAYER
GUY M. WHIPPLE — S. A. COURTIS — LEWIS M. TERMAN
HARRY A. GREENE

VOL. XI MARCH, 1925 No. 3

CONTENTS

	Page
HOW SHALL SUBTRACTION BE TAUGHT? F. B. Knight, G. M. Ruch and O. S. Lutes	157
SUCCESS, PERSONALITY, AND INTELLIGENCE W. W. Charters	169
SCIENTIFIC MEASUREMENT AND RELATED STUDIES IN CHINESE EDUCATION William A. McCall	177
A STUDY OF RE-TESTS S. C. Garrison and M. S. Robinson	190
THE EDUCATIVE VALUES OF PRACTICE TEACHING John Louis Horn	197
GRAMMATICAL KNOWLEDGE AND READING ABILITY IN FRENCH Arthur G. Bovee	204
EDITORIALS	213
REVIEWS AND ABSTRACTS	218
NEWS ITEMS AND COMMUNICATIONS	227

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1924, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus.*

W. W. CHARTERS, *University of Pittsburgh, Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

HARRY A. GREENE, *Ex-Officio, Secretary Educational Research Association, State University of Iowa, Iowa City.*

GEORGE D. STRAYER, *Teachers College, Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California.*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

MURRAY DALMAN, *Director Department of References and Research, Indianapolis, Indiana.*

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, Cal.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska.*

WILLIAM S. GRAY, *Dean, School of Education, University of Chicago.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin.*

ERNEST HORN, *University of Iowa.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean Administration, University of Minnesota.*

ROBT. H. LANE, *Assistant Superintendent of Schools Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Mo.*

ARTHUR B. MOEHLMAN, *University of Michigan and Detroit Public Schools, Ann Arbor, Mich.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools, Rochester, N. Y.*

W. J. OSBURN, *Director Educational Measurements, State Department of Public Instruction, Madison, Wis.*

J. B SEARS, *Stanford University, Stanford University, Cal.*

J. L. STENQUIST, *Director Bureau of Educational Research, Public Schools, Baltimore, Md.*

W. W. THEISEN, *Assistant Superintendent of Schools Milwaukee, Wis.*

C. W. WASHBURNE, *Superintendent of Schools, Winnetka, Ill.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XI MARCH, 1925 Number 3

HOW SHALL SUBTRACTION BE TAUGHT?

F. B. KNIGHT, G. M. RUCH, AND O. S. LUTES
State University of Iowa

There is a great deal of debate as to the relative advantages of the several ways of teaching subtraction. Textbooks of presumably equal excellence present different methods, teachers of presumably equal ability favor different methods, and writers of presumably equal competence are found in opposing camps.

The final judgment in this dispute must be based upon none other than extensive and critical experimentation. In the end it may be found that one method is as good as another. However, it is somewhat more reasonable to expect that some one method will prove to have definite, if slight, advantages over the others. Since reliable experimental evidence does not at present exist, and since *some* stand must be taken while objective evidence is being compiled in order that pupils may be taught subtraction at all, it is quite defensible to base our immediate practices upon the data at hand.

Reasoning, argument, and even theorizing should not take the place of experimentation. It is not our purpose in this paper to substitute argument for research. Our purpose is to point out that, pending scientific demonstration of the superiority of any one method over the others, those who desire to teach the *take-away* idea of subtraction in preference to the *additive* idea can do so without turning their faces away from the *known* truth. Moreover, those who wish to teach the *borrowing* technique in preference to the *carrying* or *Austrian technique* can do

so without being accused of acting in violation of any demonstrated superiority of the carrying technique.

This paper is frankly partisan rather than judicial. A partisan view is sometimes useful in that it helps to make concrete the issue involved and also to increase the demand for experimental demonstration of the facts in the case. The pages which follow present for purposes of discussion a defense for subtractive subtraction and for the borrowing technique. It goes without saying that we shall hasten to change our position if the weight of research on the issues involved goes against the theoretical case presented in the following paragraphs.

Possible Methods of Subtraction

There are two main types of subtraction, stated in terms of the ways the pupil can think out the solution. Thus, in 9 — 7, the child may think "9 less 7 equals 2," or some equivalent statement. This may be termed the *subtractive* type. We shall refer to it as "Type A."

In the same example, the child may think "7 plus what equals 9." This may be called the *additive* type. We shall refer to it as "Type B."

Again, when the pupil meets the situation

$$\begin{array}{r} 426 \\ -318 \\ \hline \end{array}$$

the 6 in units place in the minuend must in some way become 16. By the borrowing or splitting method he gets the 16 by thinking the 2 in tens place as 1 and the 6 as 16. This method we shall call "Type C."

The *Austrian* method would obtain the 16 by adding 1 to the 1 in 318. This we shall call "Type D."

For purposes of economy we shall discuss the advantages of Type A over Type B, and of Type C over Type D.

Criteria for Choosing Methods

The following criteria may be used in deciding which of several possible methods may be used in teaching a new skill. We

do not know the relative importance of these criteria, nor do we claim that they are the only ones that should be considered.

1. Which method allows the most helpful transfer from previously learned skills?
2. Which method avoids the most negative transfer or interference from skills previously learned, or to be learned subsequently?
3. Which method uses the simplest process?
4. Which method carries the most useful contributions to other aspects of the total subject?
5. Which involves the mastery of the fewest facts?
6. Which can be carried furthest into more complicated skills with reasonable ease?
7. Which can be most adequately rationalized?
8. Which admits of the most effective drill devices?
9. Which contributes best to thinking in problem solving situations?
10. Which admits of better checks?
11. Which can be best illustrated by object teaching at the beginning of the instruction?
12. Which seems to be the most natural?
13. Which is defended by experimental evidence?

Criterion 1. Most helpful transfer.—Let us discuss the matter of helpful transfer first with reference to Types A and B. When the subtraction facts are taught as additions, the previously learned facts of addition are of obviously great use; and Type B has the advantage as far as the first presentation of subtraction facts is concerned. Type A calls for the learning of facts far more independent of previously learned information than does the additive method or Type B, but we must be careful not to overestimate this first advantage. It should be borne in mind that additive subtraction, while using knowledge transferred from addition, also limits the view of the nature of subtraction. By the additive method the child first learns in addition proper that $5 + 4 = 9$. He then uses this knowledge to master the additive approach to subtraction in the situation "5 and what make 9?"

although he is thus saved the trouble of mastering the relatively independent knowledge that 9 less 5 are 4, this gain is offset by the fact that he actually runs the risk of not getting the fundamental idea of subtraction at all.

To quote from Dr. B. R. Buckingham:[1]

> One of the things I can't get around is the fact that we are calling this process subtraction. If we teach the combinations (called subtraction combinations) merely as addition combinations, it is difficult to justify either to ourselves or to the pupils the use of the term "subtraction." For example, is the question "4 and what are 9" any more obviously a subtraction question than the question "4 and 5 are what"? Moreover, when these questions are answered, the answers are identical, in both cases "4 and 5 are 9." Only by vocal emphasis is there a distinction.

There is little reason to believe that Type B has any inherent advantage, for when Type A is well taught, the additional facts used in additive subtraction are mastered through their use for purposes of variety of drill in a well-rounded attack. It is to be noted that Type B uses but one view of numbers while Type A contemplates both the additive and subtractive mode of thinking, Type A itself being the characteristic way of thinking subtraction. In reality it could be held that Type A *has* the advantage since it uses Type B in addition to itself, while Type B restricts the the child's view of numbers to but one point of view, namely, the additive.

Type C can use more transfer from previously learned knowledge than can Type D since the Austrian type of subtraction has no background at all while the borrowing or splitting type of subtraction can use certain facts concerning the changing of the form but not the value of numbers. Relative to Criterion 1 it would seem that if Types A and C when properly taught are not superior they are at least equal to Types B and D.

Criterion 2. Which avoids the most negative transfer and interference?—Here, Type A is distinctly better than Type B. Though the negative transfers of subtraction have never been pointed out the following facts appear to be true:

[1] Personal communication to the authors.

a) The additive method of subtraction is just enough like and still just enough unlike the proper methods of making change in United States money to act as a negative interference with the making of change involved in small purchases. Consider the problem: "I buy seventeen cents worth of goods and give the storekeeper a fifty-cent piece for my purchase. What change should I receive?" By the additive method the child would say: "Seven and what make ten? Two and what make five?" That is he would subtract seventeen from fifty in one fell swoop. This is just the thing he does *not* do in making change—rather, he counts up to twenty, then to twenty-five, and then to fifty. In additive subtraction and in making change he adds, but he adds in making change in an essentially different way. Here we have a case where two skills should be kept apart because they are both alike and different. Subtractive subtraction does not interfere with proper methods of making change. It properly keeps the making of change apart as a separate skill. Additive subtraction fails to do this.

b) Negative interference in additive subtraction is involved in a situation like this:

$$\begin{array}{r} 9 \\ -7 \\ \hline \end{array}$$

The child thinks "seven and two make nine." Nine is the last thing in his mind, and the last thing said or thought; but he then has to write the *two*. If this example is worked by the subtractive method the child would think "Nine less seven are two." Here he writes the number that he has last in mind.

It is possible to avoid this handicap in additive subtraction by teaching the pupil to think of the example as "nine are seven and two." In this way the pupil will indeed write the figure that he has last in mind; but probably most teachers of arithmetic will agree that additive subtraction should show great advantages in order to justify, as the normal procedure for thinking subtraction, this rather devious way of handling the numbers involved.

What effect this order would have on addition itself in work requiring the use of both addition and subtraction in close proximity is not known. A good guess would be that it would not help addition in the least.

c) Again consider the form of the same example. Additive subtraction, so far as it is influenced by the total experience with addition, is associated in the pupil's past experience with sixteen—not with two at all. Analysis of errors, especially in verbal problems, shows that this kind of negative interference exists.

d) Furthermore, the additive method is associated with the getting of something larger but the remainder is always smaller, and there is a subtle contradiction which causes confusion.

e) Consider the example:

$$\begin{array}{r} 2\tfrac{1}{4} \\ -1\tfrac{1}{4} \\ \hline \end{array}$$

Additive subtraction bids the child say: "What must I add to one-fourth to get one-fourth?" But you never do get zero in adding positive quantities. It is noteworthy that addition, per se, involves an increase. To make additive subtraction involve decrease is to inject an interference factor.

There are certain negative interferences to Austrian subtraction that should be pointed out. For example, in Austrian subtraction one always adds one. There is a real possibility that this perpetual carrying of one helps to form the frequent errors in column addition in carrying one when the carrying of some other digit is proper. Again, consider the example:

$$\begin{array}{r} 2104 \\ -1997 \\ \hline \end{array}$$

Here, according to the Austrian method, 1 is added and the 9 is changed to 10. When Austrian subtraction is used only one of the nines changes. But to the degree that the mind is set for addition previously formed habits tend to change 1997 to 2007. The difficulty involved in Austrian subtraction when there are

nines in the subtrahend is as yet an uncertain factor from the experimental point of view. The theory at least is against it.[2]

It is significant in this connection that Mead and Sears *did* find marked interference in addition after training in Austrian subtraction methods. They taught two parallel groups for four months, one by additive and one by take-away methods. During the fourth month both groups were given ten minutes daily drill in addition. The take-away group suffered little or no confusion (interference) but the additive group lost in ability so that, although they had started with a one-point lead, they ended the training period three and one half points behind the take-away group. These results are shown concretely in Table I adapted from Mead and Sears.[3]

TABLE I

Date of Tests	Take-away Group		Additive Group	
	Median	P. E.	Median	P. E.
February 9	6.7	3.5	7.7	3.6
March 9	6.5	3.1	8.5	2.6
April 6	13.4	2.7	13.6	3.3
May 4	13.4	3.3	14.7	3.6
Daily drill in addition was introduced here				
June 1	16.2	3.5	12.7	4.5

In respect, therefore, to the second criterion—which has to do with negative transfer or interference—we conclude that if the case is not in favor of Types A and C, it is at least a fifty-fifty break between them and Types B and D.

Criterion 3. Which method uses the simplest process?—We do not at present know which method is the simpler but weight

[2] It is well known that the difficulty with nines in the subtrahend when Austrian subtraction is used is offset by the zero difficulty in the minuend when the borrowing or splitting method is used. But the difficulties are of different types. The zeros in the minuend with borrowing used is not a difficulty of acting contrary to previously formed habits as is the difficulty with nines in the subtrahend when the Austrian method is used. Although we do not wish to minimize the difficulties of the borrowing method, it should in fairness be held that it is not a difficulty so far as the present criterion is concerned.

[3] Mead, Cyrus D., and Sears, Isabel. "Additive Subtraction and Multiplicative Division Tested." *Journal of Educational Psychology*, 7:261, May, 1916.

of opinion seems to be that Type B is simpler at first with the temporary advantage getting less and gradually disappearing entirely as more complicated numbers appear such as mixed numbers and compound denominate numbers.

Criterion 4. Which method carries contributions to other aspects of the total subject?—Here Type C is obviously better than Type D for the following reasons:

a) It makes dynamic use of number relations and of thinking numbers in various ways, e.g., in the example

$$\begin{array}{r} 52 \\ -17 \\ \hline \end{array}$$

fifty-two can be thought of as "forty and twelve." It further makes use of the important idea that the form of a number can be changed without changing its value; and this idea runs true in all mathematics.

b) In the subtraction of mixed numbers, Type C involves frequent thinking of such situations as this: "four and one-fourth equals three and five-fourths." This is closely allied to the processes of reducing fractions in answers, and of changing mixed numbers to improper fractions in multiplication and division. We can find no similar contributions for Type D to other aspects of the total subject.

Moreover, the subtractive method, Type A, helps to give insight into the nature of division as Type B cannot. As it is sometimes useful to think of multiplication in its relation to addition, so it is useful to think of division in its relation to subtraction. This can be done if the subtractive type of subtraction is taught. Consider the example:

$$\begin{array}{r} 27 \overline{)\,567} \\ 54 \\ \hline \end{array}$$

In explaining to the child how to finish this example, is it best for him to think, "What must I do to fifty-four in order to get fifty-six?" Or is it best for him to think "if fifty-four is taken out how much am I going to have left?"

Relative to Criterion 4, we think the argument should turn in favor of Types A and C.

Criterion 5. Which involves the mastery of the fewest facts?—The very worst that can be said under this criterion for

Types A and C is that they are on even terms with Types B and D. Since, in the end, we want 2 to be the response to the stimulus

$$\begin{array}{r} 9 \\ -7 \\ \hline \end{array}$$

whether it is taught and thought by the additive or the subtractive method, each method involves the same number of facts.

Type D of the Austrian method involves one additional process that the child has to learn—that of adding a one in the subtrahend on occasions. It is not fair to say that this fact is balanced with what has to be taught in borrowing, because the procedure of borrowing should be learned in connection with the relations of numbers in any case, whether Type C is taught or not.

Criterion 6. Which method can be carried furthest with reasonable ease?—Here the issue lies between Type C and Type D. Use the Austrian method in the example,

$$\begin{array}{r} 4 \text{ lb. } 1 \text{ oz.} \\ -2 \text{ lb. } 6 \text{ oz.} \\ \hline \end{array}$$

The child must learn to add sixteen ounces to the minuend and one pound to the subtrahend. It is true that he has to perform a somewhat similar operation in borrowing; but there his mind is set upon changing the form of a single quantity, namely the minuend. He has no such help when the Austrian method is used.

If we taught the Austrian method of division—which, by the way, is the real defense for Austrian subtraction—then Austrian subtraction should be used, not for its own sake, but for its contribution to Austrian divisions, but since very few, if any, schools in the United States use Austrian division, the real status of it under Criterion 6 is, that the further Austrian subtraction is carried into division, mixed numbers, and complex numbers, the more difficult it becomes; while the borrowing method of subtraction uses the same principle buttressed from many angles.

Criterion 7. Which can be more adequately rationalized?—Type A can be easily rationalized by having the child look to see what is left. In rationalizing the additive type, or Type B, the child has to look at objects in one place and think of them in

another. The use of objects will convince the reader that Type A is at least no harder. Type C, or the borrowing method, can be rationalized under the principle that a number can be changed in form without changing its value. As far as the writers know the rationalizing of the Austrian method or Type D is never attempted.

Of course, we readily admit that certain processes in arithmetic should not be rationalized; but we believe it is an advantage especially for the benefit of the abler pupils to rationalize a process whenever it is practicable to do so. When two methods are equal in other ways, and one method can be rationalized and the other cannot, there should be little difficulty in choosing, so far as this criterion is concerned, which method to use.

Criterion 8. Which admits the most effective drill devices?—We do not know whether effective drill devices for the "hard spots" in each of the types listed above are possible, but we do know that the splitting method, Type C, has suffered through the failure to practice certain drill devices. We suggest the following:

$$\begin{array}{r} 4\tfrac{1}{2} = 3\,\tfrac{1}{2} \\ 6\tfrac{2}{3} = 5\,\tfrac{1}{3} \end{array}$$

A few minutes of this type of drill, involving reasonable denominators and digits other than one in the numerator, removes much of the difficulty present in the subtraction of mixed numbers.

Criterion 9. Which contributes best to thinking in problem solving?—Here the issue lies between Types A and B. There are at least three situations in which subtraction is used, and in verbal problems it is well to be able to think subtraction in deciding what to do. These situations are: (a) What remains? (b) What is the difference? (c) What must I add to a known number to get another known and larger number?

In thinking out the problem: "I must work twenty examples, and I have just finished the eleventh. How many have I left to work?" the subtractive type of subtraction has distinct advantages over the additive type. The writers have noticed that some authors who teach additive subtraction have avoided the difficulty of thinking subtraction in problems which use the remainder idea by giving very few, if any, subtraction problems

where the answer is a remainder. They favor examples of this type: "I need eighty-five. I have thirty-seven. How much more must I add to thirty-seven to get eighty-five?" Of course additive subtraction fits into such verbal problems very neatly, but before an authority gives solely verbal problems that use but one of the three types of subtraction, he should make very sure that the social uses of subtraction in everyday life involve only the "How much more must I add" idea. If the reader will turn to *Thorndike's Arithmetic*, (1924 Edition,) Book 1, page 150, and figure the width of the tool house given in the drawing on that page, using in his thinking first Type A and then Type B, he will see that much could be said for the subtractive type of subtraction. A study, soon to be published, of errors made in verbal problems by children some of whom were taught the additive and others the subtractive type of subtraction reveals the fact the subtractive method avoids certain ludicrous mistakes which the additive type does not.

Criterion 10. Which admits of the better check?—Additive subtraction and Austrian subtraction admit of no check in the real sense since in checking one simply repeats the process. On the other hand subtractive subtraction and the borrowing or splitting method allow a genuine check—that of adding.

Since checking, especially from grade v upward, is far more important than is often recognized, there is some ground for the position that the uses of addition in subtraction should be reserved for checking.

Criterion 11. Which can be best illustrated by object teaching?—The issues here are so closely allied with Criterion 7 that they need not be repeated.

Criterion 12. Which method seems the most natural?—Any method that is used a thousand times by a pupil will become to him the most natural method. We are often tricked by the sheer physics of habit formation in our own thinking. There probably is, however, a certain relation between naturalness and ease of rationalization. Consider this probability in connection with rationalizing Austrian subtraction.

Criterion 13. Which is defended by experimental evidence? —The main fact to be borne in mind relative to experimental evidence is its utter paucity. Probably the best research upon this whole matter is that by Mead and Sears already referred to. It tends to favor Types A and C.

Summary

As far as the reader finds himself in agreement with the discussion above he will, by checking the conclusions of each of the criteria listed in this article, have reason to suspect that Type A and Type C have much to be said in their favor. Those who prefer to teach subtractive subtraction and the so-called borrowing method may do so without fear of overwhelming criticism from those who really understand the issues involved and without fear that the experimental evidence now available is against them. Moreover, it is to be remembered that varied amounts of motivation in teaching, if they are associated with one method of subtraction or another, will entirely disguise the true merits of the various methods.

At present choice of method in subtraction must be based on theory. All things considered the theory underlying the subtractive and the borrowing methods seems to be quite the equal of the theory underlying the additive and the Austrian methods of subtraction. It is the opinion of the writers that the theory underlying the subtractive and borrowing methods is far superior, and that pending experimental evidence it is better to practice good theory rather than questionable theory.

SUCCESS, PERSONALITY, AND INTELLIGENCE

W. W. CHARTERS
University of Pittsburgh

Mental testing is not a passing fancy; it is a permanent tool of analysis and diagnosis which will grow in power and value. It has already, in less than a decade, altered the complexion of psychological research and revolutionized the treatment of many problems of educational administration and methods of teaching. The publicity it has achieved among laymen has been amazing, and its use in technical educational fields has become universal. It has hewn out a roomy niche for itself in the halls of learning.

Every giant has his limitations, every principle has its constellation of exceptions, and no method has an unlimited field of application. Intelligence testing is no exception to this rule. Yet we find that sometimes careless administrators form special classes for students merely upon the data obtained from mental tests. Intelligence ratings are frequently placed in the hands of teachers to guide them in giving class grades. Teachers whose class marks do not correlate with the intelligence tests are occasionally placed on the defensive. Students with low intelligence quotients are sometimes considered hopelessly lost in the school world. In business also there are employment officers who accept or reject applicants upon the basis of their ability or inability to reach a critical score. These are a few of the cases which are found in sufficient numbers to give us concern and to raise the question of the limitations of the intelligence test as a basis for the prediction of success either in school or in life.

Correlation between intelligence rankings and school marks is not high enough to be considered more than prima facie evidence of its existence. The correlation runs from plus 0.40 to plus 0.60, and plus 0.50 is the accepted average. This means merely that there is a slight tendency for bright students to make high grades in class, and it means nothing more. In general, the brighter students make higher grades, but not often enough to

give us much confidence in predicting more than that they *ought* to make higher grades. That every student indicated as bright by the intelligence test will do so is never the case in any school. Some of the bright students make lower grades, and some of the less bright ones make better grades, than would be expected. As a matter of fact, it has been shown that the less bright students make on the whole better grades than would naturally be expected, while the bright students make lower grades than they ought, judged by their intelligence scores. Deductions concerning school success based merely upon intelligence rankings are wide open to criticism.

If we turn from the school field to the vocational field and seek to predict success in business from intelligence scores, the case is strikingly hopeless. On numerous occasions it has been demonstrated that the correlation between intelligence scores and success in salesmanship drops below the school correlation of plus 0.50 and hovers around zero. The bright man is no more likely to make a good salesman than is one less bright. The methods of measuring success may be faulty, but it is a peculiar fact, as well as a significant one, that no matter what the method for measuring success may be—be it earning power or judgment of executives—there is no correlation.

In the face of these facts, the question regarding the existence of other possible factors which may influence success in school life is naturally raised. To this question common sense gives us some clues. Everybody will tell you that character and personality are factors in success. Executives will volunteer the information that industry—the ability to work hard and steadily—is a potent factor in success. In some occupations social forcefulness and sociability are considered very important. The man of the street will undoubtedly tell you that the degree of interest in a vocation often spells the difference between success and failure. These are matters of common knowledge. University folk are inclined to overemphasize intellectual ability, while possibly the layman tends to underemphasize it.

Which alternative is more nearly right is no concern of ours at present. In predicting success we ought rather to explore the clues provided in the field of personality. It may be that we could better predict success if we canvassed the traits of personality. Difficult as is this task with our present technique, it must be done and the findings must be used to supplement intelligence scores. Indeed a beginning has been made in these respects. May's recent study at once comes to mind. In this significant investigation he ranked 450 Freshman students in the University of Syracuse on the honor-point basis, where *A* carries three points; *B*, two points; *C*, one; and *D*, zero. Then he ranked the students on the basis of intelligence scores, and found that the correlation was plus 0.60, which is slightly higher than the correlation usually found. He then measured their industry by a somewhat crude but ingenious method, he computed the average number of hours spent in study by each student and used this as a measure of his industry. The data secured from the students had a high degree of reliability.

The correlation between the time spent in study and honor points was plus 0.32 and shows that industry alone measured thus is not as good an index of probable success as is brightness alone. Yet when he combined industry and brightness and correlated this factor with success as measured by honor points, the correlation rose to plus 0.80. Intelligence alone is plus 0.60; industry alone is plus 0.32; intelligence and industry combined is plus 0.80. Thus we see in this typical study what everybody expects—that if the student is both bright and industrious, his chances for making high grades are appreciably increased.

In salesmanship, as one type of business, other traits of personality show interesting correlations, particularly in view of the fact that the correlation of mental alertness with success in selling is approximately zero. We see, for instance, on the basis of the Carnegie adaptations of the Downey character tests, that such a factor as coordination of impulses or the ability to keep two things in mind at once in handling a situation has a correlation of plus 0.68, which is higher than May's correlation of plus 0.32 between

the honor points and industry or the correlation of plus 0.60 between honor points and intelligence. Such another factor as "always working at the highest speed" has a correlation of plus 0.46, and these two combined provide us with a correlation of plus 0.76 which is almost as high as May's figures on brightness plus industry in study.

If it were necessary to substantiate the opinions of common sense—that personality is a major factor in success—other cases could be cited. Even though the facts are meager, because as yet no great advance has been made in the measurement and analysis of the traits of personality, it is confidently to be expected that eventually methods of measurement will be provided and that, when these have been secured, we shall in all probability be able to predict success to a much closer degree than at present. That a perfect prediction can ever be made is not likely, there are too many variable factors. The degree of the trait may vary from year to year. For instance, the lazy boy of sixteen may be a glutton for work at twenty-three because he is older and is married, or has found an interesting job. The girl whose health is superior on the day of her tests at sixteen may become an invalid before she is eighteen. In the final stages when the technique of trait measurement has been perfected, the prediction cannot be perfect. It will merely be closer than is now possible. Mental tests are exceedingly important in measuring mental alertness, but they have serious limitations as a measure of the total ability of an individual.

If this be true, they must be handled with extreme care so that the students may not suffer. That, under carelessly handled conditions, students do now suffer is currently known by all. In using intelligence tests, therefore, a few precautions must be taken. These I shall take the time to enumerate, since intelligence tests have been widely adopted as a routine implement in the crafts of teaching and administration.

Information about intelligence scores should only be given with extreme care. There is altogether too much casual discussion of upper and lower quartiles, brightness and dullness, and

so on in the presence of the pupils discussed. So frequently as to amount to cruelty, a teacher will discuss the student, in his presence, with someone else and speak of him as being slow or dull, or as having a low intelligence quotient and with a complete lack of regard for the feelings of the child. We know from the studies which have been made that knowledge of their mental rating causes discouragement on the part of ten percent of the students. Boys and girls have been crushed by the information that they were below the average in mentality. They have given up chosen careers, perhaps unnecessarily, have experienced continued periods of depression, and have lost ambition to succeed.

Such procedure is worse than cruel—it is false and stupid. What right has any person to predict failure because an intelligence quotient is somewhat low? To be sure, three percent of the pupils in the school may be so defective as to be hopeless, but the other ninety-seven percent may possibly be able to develop compensating traits of personality which far outweigh mediocrity in mental ability. We of the classroom are inclined to feel sorry for the industrious child with a comparatively low mental score, who by dint of hard work keeps abreast of his class. "Isn't it unfortunate," we say, "that he isn't bright?" We should rather say, "Isn't it fortunate that he is so industrious, ambitious, and friendly? What a fine chance he has for success in the game of life where there is no particular correlation between success and brightness." It is high time for all of us to follow the wise observers of men and affairs and to give to personality the credit and place of importance it deserves. In making an inventory of the assets of a man industry, forcefulness, leadership, sympathy, and ambition are major headings, which combined yield nothing in importance to brains and mental brightness.

Such being the case I repeat that extreme care should be taken in letting children know their mental scores. This care should be both positive and negative. Negatively the scores should be kept from them except where it is extremely necessary to make them known. Positively, they should always be accompanied by a supplementary statement about compensatory traits.

Four cases which require different treatment at once appear. If a student is bright and has fine traits of personality and character, he is likely to be so successful that he will naturally feel comfortable and satisfied with his progress. It is unnecessary for him to be told how bright he is. If he is bright and has some fine traits but lacks confidence to the extent of being chronically depressed, it may be necessary to reveal his strong traits in order to establish the necessary confidence. If the student is bright but lacking in the traits of industry and ambition, his intelligence score may be revealed to him, but only with an emphatic statement to the effect that he must develop his weak qualities so as to do what is expected of him. If a student is not bright mentally and learns of the fact, we may be sure that he will be discouraged by the information. So again if it is necessary to let him know his score, or if he inadvertently learns it, the teacher must hasten to explain to him that he has compensating qualities or may develop them, that success is quite as dependent upon other things as upon brightness, and that it is better to be mediocre in mentality and strong in industry and ambition than to be bright but lazy and shiftless. Too much attention cannot be given to this. Because of the thoughtless remarks of some teachers, children often develop inferiority complexes which they are never able to outgrow. Success is too complexly conditioned to be handled by any formula so simple as mental alertness and intellectual dullness.

In the second place, classes should be only tentatively sectioned upon the basis of intelligence tests. The sections should not be permanent in their composition. This does not mean that the lazy, but superior child should be placed in a lower section, when he fails to maintain his ranking; he needs to be kept where he is and the pressure increased. Yet the highly nervous child or the sickly pupil of superior mentality may be benefited by a change to what we call a lower section. And particularly may the child of great industry and ambition, who according to the intelligence ratings belongs in the next lower section, be advanced to a higher section when the teacher feels that he can successfully carry the

work because of his industry. The door should not be shut in the face of ambitious youths, just because there is a slight question as to their mental level. As I have said, success is far too complexly conditioned to be measured by a formula which involves mentality only.

In the third place, we should have records not only of intelligence quotients and of educational marks, but of personality quotients as well. The school record of a child should contain not only his school grades and his mental score, but his rating as well in certain fundamental traits of personality such as industry, ambition, friendliness, dependability, initiative, self-confidence, and resourcefulness. Only by these means can we hope approximately to predict his performance in succeeding grades or in his later adult life, and only from such data can we derive an index of his personal accomplishments. It may be more significant that a student raise a rating of *D* in self-confidence to a *B* than that he be the brightest student in his class. If somewhere along the years of his schooling he suddenly rises in his rating in ambition, the prosy symbol on his record may mark the dramatic turning point in his life. A low, but improved—or gradually improving—grade in initiative may often be a mark of greater progress than a superior grade in intelligence or a grade of *A* in class work. The traits of personality which we often treat so casually in our records need to be brought out into the focus of pedagogical attention and given the position they deserve.

I have seen many bright executives fail because they were irritable and unfriendly, and conversely I have witnessed an immediate rise in efficiency when after effort they began to overcome these faults. I have also seen men of superior intelligence and training completely and despondently fail because of lack of self-confidence, and have again witnessed a consistent and spectacular rise of these same men when their confidence in themselves was restored. Men and women who have become nervous and irritable because they did not know how to relax have prolonged their vocational careers and conspicuously increased their happiness by the mere development of the ability to take things easy.

In the presence of all these facts it is self-evident that ideals and traits play a considerable part in all measurements of people. Their's is a supplementary rôle; they compensate intelligence. Success cannot be predicted with safety by tests of mental alertness alone. Mental and scholastic records must be supplemented by ratings of personality traits. Teachers and administrators can use intelligence quotients safely within definite limits, but again it must be remembered that success is too complexly conditioned to be predicted by a single mental formula.

SCIENTIFIC MEASUREMENT AND RELATED STUDIES IN CHINESE EDUCATION

W. A. McCall
Teachers College, Columbia University

[*Continued*]

One of my tasks was to travel about the country lecturing at various universities and to educational institutes and associations. The reader may be interested in the technique of such lectures. An account of the first of my six lectures at the Tsinan Institute will give the reader an idea of the elaborate courtesy which attended all such occasions.

When I arrived at Tsinan Institute with my interpreter, Dr. T. T. Lew, we were met at the curb by the four leading educators of the Province. All four bowed and shook their own hands. We bowed, shook our own hands, descended from the rickshas, and bowed again. They bowed us to precede, and we bowed them to precede, and they finally pushed us through the moon-shaped gate in the wall into the first court. They bowed us to precede, and we bowed them to precede, and they finally pushed us through the second gate into the second court. They bowed us to precede, and we bowed them to precede, and they finally pushed us through a door into a covered hall with a reception room on each side. They bowed us into the reception room, and bowed me to the seat of honor. I protested my youth. It was offered to Dr. Lew, who was also a guest. It was offered to each in turn, and each protested his unworthiness. With the assistance of Dr. Lew I was finally prevailed upon to take the chair. A similar contest in courtesy took place over all the other seats. Then tea and cakes were served.

They bowed us out of the reception room into the covered hall and into the third court, and through the final door into the lecture hall, where the audience patiently awaited our coming. They led the way to the lecture platform. One bowed as he introduced us, and we bowed in return. The audience rose in welcome, as

the pupils do to their teachers at the beginning of each school day. My interpreter and I bowed in unison to the audience. The audience bowed in return and sat down. I began my address, and, as I did so, glanced at my tiny, frail, stooped, and utterly weary-looking interpreter. I finished the first unit of thought, and turned to him for help. A miracle happened. Something electric passed through him. He rose on his tiptoes, his immobile oriental face lightened, his eyes flashed with the humor or the fire of his words, and every particle of his being became gracefully articulate. A powerful, confident voice rose and fell in perfect accord with his ideas. He was not translating but interpreting, as he clothed my prosaic ideas with poetry. He was a natural orator, and the master of our audience to the last word. Then, as if exhausted by the effort, he subsided once more to frailty.

At the end of the lecture we bowed in unison and the audience rose, bowed, and remained standing until we were escorted from the room. At the first door, we turned back and bowed our hosts to escort us no farther, although certain that our efforts would be fruitless. At the next door and next gate and last gate we turned back and bowed farewell, attempting by physical force to prevent further trouble to our hosts. They continued against our protest, reserving their farewell bow until they had reached the curb, where they had awaited our coming. They continued bowing until our rickshas carried us around a curve in the street. This process was duplicated exactly at each of the six lectures we gave at the Tsinan Institute, and was the ordinary, conventional courtesy for such occasions, but the etiquette was reversed, when a delegation paid an official call of welcome at our hotel.

The most difficult lecturing situation which I faced was at the Institute on Scientific Supervision attended by over three hundred delegates from all the provinces of China. This Institute was called to plan the national test survey for the second year, and to train the delegates in the practices involved in intelligence tests, educational tests, practice tests, and other scientific instruments created during the year. I spoke in English as rapidly as I could

talk, then three interpreters began at once, interpreting what I had said into the three chief dialects represented among the delegates. Among the delegates were individuals who knew two or more dialects. They aided by tossing words about among their neighbors. Each group in the audience listened to its own interpreter with surprisingly little confusion.

Besides lecturing and giving technical assistance to those engaged in devising the instruments of measurement, one of my duties was to help plan, organize, and conduct demonstrations of testing and to make recommendations based thereon. Sometimes the chief difficulties were connected with the transportation of the testing staff and observers. An illustration of this difficulty is afforded by a survey conducted just before I left for America.

The delegates of the Institute on Scientific Supervision accepted the invitation of a former Premier of China to test the nine hundred children in his famous orphanage in the Western Hills, about nine miles from Peking. The city of Peking and interested citizens contributed many automobiles. In order to reserve these for the fifty examiners, who needed to arrive early, and in order to manifest my predilection for democracy, I elected to ride in a common ricksha as did the other two hundred and fifty delegates. As we were passing through acres of pink and white lotus blossoms, the first contingent of examiners roared by in every make of automobile from the police auto for criminals through the latest limousines to a Red Cross ambulance loaded with lady examiners.

When the head of the orphanage heard I was coming in a lowly ricksha, he sent back an auto for me. Courtesy required me to transfer to the car, which proceeded at a fearful pace, the siren blowing continuously to within a mile of the institution, where a tire gave way. Unable to wait for the necessary repairs, I mounted a small donkey and held up my feet to keep them from dragging. Thus nobly mounted, and with the owner whipping his lazy animal from the rear, I made a truly Imperial entry through the gate of what was once the palace of a Manchu prince,

to the great amusement of the assembled multitude and the glory of Western science!

The next step in this presentation should be an evaluation of the projects, particularly of the tests. If this evaluation were to be a relative one, it would consist principally in comparing each test with its nearest equivalent in the United States, because almost every type of intelligence or educational test has reached its highest development in America. The tests constructed by my Chinese and Mission colleagues, when checked against this standard, merit, I believe, the conclusion that in almost every case they are equal to, and in most cases they are superior to, the tests in America which served as suggestive models. The seventy-nine processes described in chapter v of *How to Experiment in Education* were used as guides in constructing the tests for China. Anyone familiar with how tests are constructed in the United States can make a rough check of my conclusion for himself by reading these processes.

Any attempts to justify my evaluation would require much technical discussion. Here only a few illustrations can be given. The non-language intelligence tests with the exception of one which is a novel and promising experiment, were constructed in harmony with Herman Liu's Doctor's dissertation, which is an elaborate study of parts of the previously existing non-language intelligence tests. Improvements in the verbal intelligence tests are of the following sorts: (1) more emphasis upon difficulty and less upon speed, (2) the simplification of scoring, (3) the mechanical arrangement of material so as to facilitate stencil scoring, (4) the location of preliminary tests so as to save printing costs and to prevent pupils working on the regular test during explanation and preliminary testing, and (5) the utilization of recent critical studies of various tests.

An improvement in the Revision of the Binet-Simon Scale is the use of a special range-finding test for determining more objectively than heretofore the equivalent of a pupil's basal age. The silent-reading tests, both Chinese and English, are so constructed that they may be scored more objectively than usual,

and most of them make use of a device which permits the test papers to be used over and over again, thus effecting a great saving in cost. Furthermore, their content was based on Chen's and Thorndike's word-counts of the Chinese and English languages respectively. One unique superiority of the composition and handwriting scales is that they give practical recognition, for the first time, to the fact that the scale values on such scales when used in the past were merely for scoring purposes. The scales, as used in China, have values comparable with any performance scale. The superiorities in the realm of physical measurements are greater still. Thus the score card and point system for standard athletic games is incomparably the best work of its kind. All mental tests have one marked superiority. All, on each level, are scaled in such a way as to permit automatic interpretation of results, and so as to secure comparability of scores from test to test.

The outstanding inferiority of tests in China is inadequate standardization. This is a defect which time and surveys already under way and others which are planned for the future will correct. Even so, there are not a large number of tests in America whose published norms are based upon more than two thousand pupils. In the case of the linguistic Chinese tests, there may be other defects invisible to me, because I do not know the Chinese language. To be illiterate in Chinese does not, however, exclude me from a fairly clear knowledge of the content of these tests, for all tests were translated for me two or more times. Furthermore, the statistical technique for evaluating the content of tests has become so refined that many times I was able to locate the linguistic defects invisible to those who knew Chinese but not the statistical procedure.

To the charge of possible inadequate standardization we must add the charge of imperfect standardization. In many of the schools we could not maintain standard testing conditions to the extent possible in the United States. We were frequently forced to choose between the two evils of employing inadequately trained local examiners or well-trained examiners brought from the

university—say a hundred miles away—whose dialect was scarcely understood by the pupils. Again, difficulties were caused by the lack of desks in some schools or two or more pupils at the same desk in others, for the double-desk system is widely used in China. I went into one room where I felt the examiner was not moving about the room enough to see whether the pupils were working on the proper page of the test at the proper time. The examiner told me that he had tried to follow my instructions but that the floor planks flew up when he stepped on them and disturbed the pupils. I went into the next room, and told the examiner there to see that the pupils were working on the *right* page. He hastened to direct the pupils to leave the *left* page on which they were working and should be working. This taught me that I also had a dialect, for I had meant the proper page instead of the right-hand page.

Correct age records for the pupils were extremely difficult to secure. Letters often had to be dispatched to the parents for this information. And then there are many ways of computing age in China. In one school system a circular had been sent to all the principals stating the standard chronological age for each grade. Before our arrival, the pupils had been instructed to place the standard age for the grade on their test blanks. The school records had also been changed, accordingly. This was not evidence of any antagonism, but of a human desire to appear well to "company." As the survey of the schools of Soochow began, there was a sudden drop in temperature. It was necessary to defer the testing until noon to give the bright sun a chance to thaw out the fingers of the pupils. There is no heat in most Chinese schools even in the coldest days of winter. To think that standardization difficulties are peculiar to China would be a mistake, but they are certainly more numerous there than in the United States.

The possibility of one other defect, particularly in the case of intelligence tests, must be recognized. The extensive use made of American intelligence tests as models is justifiable only in case the general nature of intelligence in the United States is essen-

tially the same as the general nature of intelligence in China. All competent psychologists with whom I have consulted believe that the nature of intelligence is the same in all races. They believe, further, that a test of the ability to perceive relationships among objects and ideas for example is just as symptomatic of intelligence in China as it is in America.

As will be evident from an inspection of the projects of the year, the purpose was not to study numerous interesting psychological problems. The purpose was rather the more tedious one of creating those instruments indispensable to such studies. Incidental to the scaling and standardization of the tests it was possible to collect data which are now being so treated as to yield important information. All that will be given here is a brief description of a few of the inventions or discoveries which have more or less universal validity and which are thus likely to be of interest to American students. These are:

1. *A simple, generalized formula for allowing for chance regardless of the number of alternatives.*—This formula will prevent the pupil who lacks ability but who marks rapidly and at random from making a higher score than the pupil who is slow or cautious but sure. In the past, it has been customary to ignore this error when the alternatives were more than two. (Prior publication by others)

2. *A generalized formula for the inclusion of elements in alternatives such that full knowledge will be required before correct answers can be given.*—Thus, in a silent-reading test, a pupil may be asked to read a passage and then mark the one of five drawings which is described by the passage. The problem solved by this formula is how to distribute the difficult elements of the passage among these five drawings so that a chance knowledge of the meaning of one or a few words will not give the clew to the correct alternative, or will not so restrict the range of choice as to invalidate the formula for allowing for chance.

3. *A new method for constructing product scales such as composition and handwriting scales.*—This method is far less laborious than the standard method in use, and it yields directly both a product-scoring instrument and a product scale with values comparable with those for ordinary performance tests.

4. *A method for the construction of a C-scale.*—The C-score yielded by this scale gives an automatic reading of a pupil's proper classification in a standard school.

5. *A straight-line equation for a C-scale, and an identical equation for all C-scales.*—One of the two most surprising discoveries to date, is that the C-scale curves or grade-progress curves for all tests may be

superimposed on each other. This means that after determining the T-scale for all tests and the C-scale or grade norms for one, it would have been possible to use this one C-scale for all the other tests. An identical table for computing C-scores is being used for all tests on each level. The C-scale curves for the three scaling levels, namely grades I-IV, IV-VIII, and VIII-XII, have a different slope, but all are straight lines. Whether this C-scale equation will fit the tests constructed during the second half-year remains to be seen.

6. *A simple and identical logarithmic equation for all B-scales.*—This is the other surprising discovery. The identity of the B-scale curves or age-progress curves for the different tests was less close than in the case of the C-scale, but it was close enough to justify the use of a common table for computing B-scores. It has since been shown that this equation holds for the United States also.

7. *A new method for the construction of intelligence tests.*—The chief merits of the method are simplicity and flexibility. It permits a content of very great variety, and the testing of a large number of different mental traits. It may be used in the construction of tests for all intelligence levels from the first grade through the university. The same uniform, simple, controlled response is required throughout the entire test and irrespective of the level of intelligence being tested. Few preliminary directions are required, and there are no directions after the pupils begin to take the test. This makes the application of the test a very simple matter. Finally, the method effectively overcomes the great difficulty of constructing test elements which are sufficiently difficult to measure the maximum mental subtlety of gifted university students.

8. *A new cumulative-record card and record system for pupils.*—This scheme substitutes a modern system of T-C-B-F units for the obsolete A-B-C-D-E or percentage system in use in both China and America. This system yields for any examination given by a teacher all the interpretative units usually found only in connection with standardized and scaled tests. Appropriate cards and filing cases needed for the use of this system are being prepared by the Association.

9. *A new and unusually simple formula for measuring the efficiency of a school and school system.*—This formula utilizes direct rather than correlated efficiency indices. The formula operates automatically to discourage school systems from having an undue percent of age retardation. It prevents a school from creating the appearance of efficiency by refusing to graduate students at the normal time, and from crediting to itself a superior efficiency which is really due to location in a neighborhood which contributes gifted children to the school. The formula operates automatically to encourage the community to lengthen the school term to the optimum length, and to secure steady attendance on the part of the pupils. It operates to reduce the cost of school instruction to the lowest point. It strives not to cripple the effectiveness of the instruction, to effect unfavorably the morale of the teachers, or to drive the more efficient teachers from the profession. The formula gives an auto-

matic index of the extent to which the staff of the school system is entitled to higher salaries because of superior efficiency. It thus operates as a powerful motivator to the harmonious cooperation of all the agencies of the school for increasing the efficiency index so as thereby to justify a salary increase.

My reception by official China and my farewell by official Japan is symbolic of the kindliness I found. When I boarded the train at Nanking for my first trip to Peking, I discovered that a guest ticket had been provided, and that the trainmen had been told of my coming, in order that they might make my long journey as comfortable as possible. When I went to the Japanese Embassy in Peking to get a visè on my passport, I was handed a letter. For all I knew its strange symbols said: "This is a dangerous character, watch him." I tried its potency at the much-dreaded Japanese customs. My baggage was passed uninspected. I tried it again in Kyoto, and a guide was provided to show me the Imperial Palaces and Gardens. When the earthquake came in Tokio, I wished for it mightily for a few minutes, but it was at the Imperial Hotel, the only one which was spared!

When I laid an ambitious program for the year before Director Tao and asked whether the budget could carry it, I was told to be limited only by the technical assistance available in the nation and the time at my disposal. As I anticipated, the funds were exhausted before the year was half over. Director Tao disappeared for a few days and came back with ample funds for the remainder of the year. As I understand it, the Chinese Cabinet voted some of it, and General Chi, who at the time of this writing is fighting for the control of Shanghai, contributed a part.

The reception given the movement itself is even more indicative. In the middle of the summer, representatives of the Chinese Christian Educational Association met on top of the Kuling Mountain and voted unanimously for a test survey of the mission schools of China, with the double purpose of evaluation and supervision. Two weeks later, three hundred Chinese delegates from all the important provinces and cities of China, met in Peking and voted for a national test survey of the public schools.

They remained two weeks for an Institute on Scientific Supervision to learn the technique of using the instruments that had been created during the year and to receive training that they might assume local control of the survey in their respective cities under the general direction of the Chinese National Association for the Advancement of Education and its technical staff, composed of those who had worked most intimately with me during the year. I had expected to remain in China a second year in order to participate in the proposed survey, but at this point I was called back to America. Consequently, the year's survey of all the public and mission schools in the twenty-three largest cities and many surrounding smaller cities of China, was carried on without the constraint of that unfailing courtesy given to guests.

The survey was planned to move in a wave through China. It was financed locally; all preliminary arrangements were made by local directors, in accordance with preliminary printed directions sent. The plan called for the completion of a survey in each large city within a week. When the staff of the Association arrived on Monday morning, hundreds of normal-school students were waiting to receive training in the technique of applying and scoring the survey tests, and of tabulating and computing the results. Two or three days were given to training, demonstrations, and practice. On Wednesday a third of the schools were tested and by midnight that night all tests were scored, tabulated, summarized, and each step checked. By Friday night the survey was completed and copies of the detailed and summarized results were ready for the series of teachers' meetings to be held on Saturday, where the results were interpreted and ways of using them suggested. Then the survey staff moved on to the next city.

Since the movement has been vigorously launched and favorably sponsored by the leaders in Chinese education, the next need, as I see it, is for creation and publicity to give way to application in teaching, supervision, and experimentation, and for criticism to be developed. The last mail from China brings me three pleasant items of information. Some critical papers ap-

peared on the program at the last national gathering of those with whom I worked. The Board of Director of the Association have also elected me to a sort of honorary advisory position which will give me an excuse for maintaining an active contact with the movement. And a book in English, published by The Commercial Press, Shanghai, gives an account of the national test survey, together with a summary of the results found.

I cannot close this already over-long article without a word of appreciation to those who have made the year's program possible. I count it a very rare privilege to have been associated with the Chinese National Association for the Advancement of Education in the year when it inaugurated its nation-wide program of test construction on a scale never before attempted, and enlisted for this work some of the best-trained minds in Chinese education. For this gave me an opportunity to witness in these educators of China such a spirit of generous cooperation and such a devotion in spite of the drudgery inseparable from careful research as is without parallel in my experience.

LIST OF TESTS CONSTRUCTED AND PROJECTS UNDERTAKEN BY THE MEMBERS OF THE CHINESE NATIONAL ASSOCIATION FOR THE ADVANCEMENT OF EDUCATION

I. Intelligence Tests

1. Butler Non-Language Intelligence Test, Forms 1 and 2. For grades I-IV. By Alice Butler, professor of psychology, Ginling College for Women.

2. Liu-Bradshaw Non-Language Intelligence Test, Forms 1 and 2. For grades IV-VIII. By Dr. H. C. E. Liu, educational secretary, Chinese National Y. M. C. A., and Annie Bradshaw, professor of psychology, Laura Haygood Normal School.

3. Liao Verbal Intelligence Test. Scale A, Form 1, for grades IV-VIII; Scale B, Form I, for grades IV-VIII. By Dr. S. C. Liao, professor of education and principal of Model Secondary School, National Southeastern University.

4. Lew Verbal Intelligence Test, Forms 1 and 2. For grades VIII-XII. By Dr. T. T. Lew, dean of College of Theology, Yenching University, and professor of psychology, Peking National University and Peking Teachers College.

5. Chang Verbal Intelligence Test. Scale A, Forms 1 and 2; Scale B, Forms 1 and 2; and Scale C, Forms 1 and 2. For college and college entrance. By Professor Y. C. Chang, professor and head of department of psychology, Peking Teachers College.

6. Luh Revision of the Binet-Simon Intelligence Scale, Form 1. For ages 3 to adult. By Dr. C. W. Luh, professor and head of Department of Psychology, National Southeastern University.

7. Terman Mechanical Intelligence Test, Forms 1 and 2. For grades V-XII. By E. L. Terman, professor of education, Yenching University.

II. Educational Tests

8. Chen Silent Reading Test. Forms 1 and 2, for grades I-IV; Forms 1, 2, 3, 4, and 5, for grades IV-VIII; Forms 1 and 2, for grades VIII-XII. By H. C. Chen, professor of psychology, National Southeastern University.

9. Westbrook English Silent Reading Test, Forms 1 and 2. For grades V-VII. By Dr. C. H. Westbrook, dean of the School of Education, Shanghai College.

10. Brinkley English Silent Reading Test, Forms 1 and 2. For grades V-XII. By S. L. Brinkley, professor of education and English, Soochow University.

11. Gage English Vocabulary-Reading Test, Forms 1 and 2. For grades V-XII. By Brownell Gage, dean of Yale-in-China.

12. Keys English Vocabulary-Opposites Test, Forms 1, 2, 3, and 4. For grades V-XII. By J. N. Keys, professor of education, Canton Christian College and supervisor of English, Kwantung Province.

13. Chang National Language Auditory Comprehension Test, Forms 1 and 2. For grades II-VII. By S. Y. Chang, professor of English, National Southeastern University.

14. Keys English Auditory Comprehension Test, Forms 1 and 2. For grades V-XII. By Professor J. N. Keys.

15. Anderson Comprehensive English Test, Form 1. For grades VIII-XII. By E. J. Anderson, professor of psychology, Shanghai College.

16. Keys English Grammar-Idiom Test, Forms 1 and 2. For grades V-XII. By Professor J. N. Keys.

17. Chou Mixed Classical and National Language Composition Scale, Form 1. For grades I-XII. By Dr. H. C. Chou, professor of education, Amoy University.

18. Yui National Language Composition Scale, Form 1. For grades I-XII. By T. Y. Yui, professor of education and principal of Model Elementary School, National Southeastern University.

19. Bullock English Composition Scale, Form 1. For grades VI-XII. By A. A. Bullock, professor of education, Union Normal School, Wuchang.

20. Yui Formal Handwriting Scale, Form 1. For grades I-XII. By T. Y. Yui.

21. Yui Running Handwriting Scale, Form 1. For grades IV-VIII. By Professor T. Y. Yui.

22. Chen Spelling Scale, Forms 1 and 2. For grades III-VIII. By Professor H. C. Chen.

23. Terman Mixed Fundamentals of Arithmetic, Forms 1 and 2. For grades III-VIII. By Professor E. L. Terman.

24. Yui Arithmetic Problems Test, Forms 1, 2, 3, and 4. For grades III-VIII. By Professor T. Y. Yui.

25. Chang Algebra Test. Fundamentals, Forms 1 and 2, for grades VIII-XII; Equations, Forms 1 and 2, for grades VIII-XII; Problems, Forms 1 and 2, for grades VIII-XII. By C. Y. Chang, president of the Paotingfu Normal School.

26. Yui Practice Tests in Integers and Decimals, Forms 1 and 2. Fifty-eight practice lessons for grades IV-VIII. By Professor T. Y. Yui.

27. Chen General Science Test, Forms 1 and 2. For grades III-VIII. By Professor H. C. Chen.

28. Yeung Chinese Geography Test, Forms 1 and 2. For grades IV-VIII. By K. T. Yeung, professor of psychology, Canton Christian College.

29. Chen Science Test, Forms 1 and 2. For grades VIII-XII. By Y. G. Chen, dean of the Department of Physics and Chemistry, Peking Teachers College.

30. Tsu Chinese History Test, Forms 1 and 2. For grades VI-XII. By T. L. Tsu, professor of history, National Southeastern University.

31. Lewis Household Arts Objectives and Test, Scale A, Form 1; Scale B, Form 1; and Scale C, Form 1. For grades VI-XII. By Ida Belle Lewis, educational secretary for the Methodist Woman's Board of Missions for China.

32. Chang-Ling Citizenship Objectives and Test. For grades I-XII. By Dr. P. C. Chang, director of Secondary Education Division of the National Association for the Advancement of Education, and Dr. P. Ling, dean of Nankai College.

33. McCloy Health Test. For grades I-XII. By C. H. McCloy, professor of physical education, National Southeastern University, and director of physical education, Kiangsu Province.

III. Special Tests and Related Projects

34. Terman Survey Non-Language Intelligence Test, Forms 1 and 2. For grades IV-VIII. By E. L. Terman.

35. Cha Survey Battery of Educational Tests, Forms 1 and 2. For grades IV-VIII. By L. C. Cha, dean of Peking Teachers College.

36. McCloy Score Card and Point System for Standard Athletic Games of Skill. For grades I-XVI. By C. H. McCloy.

37. McCloy Curriculum of Objectives in Physical Education. For grades I-XVI. By C. H. McCloy.

38. Chu-Tu School-Building Score Card. By Dr. J. P. Chu, professor of statistical methods, National Southeastern University and dean of the First Girls' Normal School of Kiangsu, and Y. C. Tu, professor of civil engineering, National Southeastern University.

39. Chen Character Count of the National Language. By Professor H. C. Chen.

40. Chen Idea Count of the National Language. By Professor H. C. Chen.

41. Chen Score Card for the Difficulty of Words, Sentences, and Paragraphs. By Professor H. C. Chen.

42. Chu Standardization of Measurement and Statistical Terminology. By Dr. J. P. Chu.

43. Chuang Standardization of Psychological Terminology. By Dr. C. H. Chuang, professor of psychology, Tsing Hua College.

44. Cheng-Meng-Tao Standardization of Educational Terminology. By T. H. Cheng, professor of education, National Southeastern University, Dr. H. C. Meng, professor of the philosophy of education, National Southeastern University, and W. T. Tao, director of the National Association for the Advancement of Education and dean of the School of Education, National Southeastern University.

45. Hsüeh Cumulative Record Card for Standard Tests. By H. T. Hsüeh, statistician, Chinese National Association for the Advancement of Education.

46. Chen-Chu Survey of the Hangchow School System. By Professors H. C. Chen and J. P. Chu.

47. Liao Experiment with the Dalton Plan in a Chinese Secondary School. By Dr. S. C. Liao.

48. Wells Slide Rule for Converting Chinese Old to Chinese New Age. By Ralph C. Wells, principal of the Point Breeze Academy of Weihsien.

49. McCall Permanent Exhibit of Tests and Standards in the Temple of Emperors, Peking. By W. A. McCall and others.

A STUDY OF RE-TESTS

S. C. Garrison and M. S. Robinson
George Peabody College for Teachers

The purpose of this article is to present data on re-tests when the National Intelligence Test is used and to compare these with similar scores secured through the use of the Stanford Revision of Binet-Simon Test. In a previous article one of the writers presented the results of re-tests by means of the Stanford Revision[1] to which the reader may refer if interested in a detailed discussion of that subject.

In judging the practical worth of a test several factors must be considered: (1) The correlation with reliable criteria should give fairly high coefficients. (2) When repeated the test should give results similar to those secured at the first testing. (3) The test should be easy to administer.

The National Intelligence Test was given to the pupils of grades III to VIII inclusive, at the Peabody Demonstration School during the first week of April, 1922; the second week of January, 1923; and the last week of September, 1923. Form A1 was used for the first two testings and A2, for the third. We have three records for 131 pupils and two for 140.

The Stanford Revision of the Binet-Simon Test is given to the pupils in the demonstration school each year. We have now the results of four or more testings of the pupils used in this study.

Conditions

The pupils of the school are given standardized educational tests at different intervals during the year. No pupil used in the study had taken less than four group intelligence tests before the first testing with the National Intelligence Test, and they had also taken a number of educational tests. These pupils had acquired whatever advantage may be gained by practice in taking

[1] Garrison, S. C. "Additional Re-tests by Means of the Stanford Revision of the Binet-Simon Test," *Journal of Educational Psychology*, 13:307-12, May, 1922.

tests previous to the time when the National Intelligence Test was given.

With very few exceptions these pupils try to do their best when taking tests. It is the practice here to tabulate results of the educational tests and post them in the schoolroom where all the students, parents, teachers, and anyone else interested can examine them. The children know that the results of the intelligence tests are kept and consulted by their parents, the teachers, and the school principal.

As a rule, the work done in this school is of a superior type. Because of a good tuition fee, which acts as a selective factor, the pupils come from the best homes of the city. The median intelligence quotient (Binet-Simon) for the children in the first eight grades is 112.

The Data

In tabulating ages, we have listed as ten years old all pupils who have passed their tenth birthday but have not yet reached their eleventh. The age given in the tables is in every case that of the child at the last testing. The tests were all administered and graded by the writers or by graduate students in educational psychology. All the persons who did any testing had had extensive preparation and practice in giving tests. The persons, however, who did the re-testing with the Binet-Simon Scale were ignorant of the previous records made by the pupils.

The distribution of the differences in intelligence quotients found by re-tests with the National Intelligence Test ranged from —18 to more than +20. From Table I it appears that approximately 50 percent of the differences between the intelligence quotients of Test 1 and Test 2 lies between +6 and —6, inclusive, while those for Test 1 and Test 3 fall between +7 and —7. For the ten-months interval, the data secured from comparison of the second and third tests, similar limits are +5 and —5. When the Binet Test was used we found the limits for the re-tests after an interval of one year to be +4 and —2; for a two-year interval, +4 and —3; and +3 and —2 for the second and third tests. An inspection of Table II shows that in the National

Test 55 percent of the pupils made higher intelligence quotients where re-tested, 42 percent made lower, and 3 percent remained the same. For the Binet-Simon data corresponding figures are respectively 55, 38, and 7 percent.

Our data seem to show that the National has a tendency to scatter the pupils over a wider range than the Stanford Revision.

TABLE I. LIMITS INCLUDING APPROXIMATELY 50 PERCENT OF THE DIFFERENCES IN INTELLIGENCE QUOTIENTS FOUND BY RE-TESTS WITH THE NATIONAL INTELLIGENCE TEST

Testings	Binet-Simon	National
First and second tests—10-months interval	+4 and −2	+6 and −6
First and third tests—20-months interval	+4 and −3	+7 and −7
Second and third tests—10-months interval	+3 and −2	+5 and −5

It also gives more variable results in re-tests, and the intelligence quotients tend to increase faster. We found an average gain of

TABLE II. PERCENT OF PUPILS WHOSE INTELLIGENCE QUOTIENTS INCREASED, DECREASED, OR REMAINED CONSTANT WHEN RE-TESTED

Percent	Binet-Simon	National
Increase	55	55
Decrease	38	42
Constant	7	3

1.7 for the Stanford Revision over a two-year interval, while that for the National over a twenty-months interval was 3.7. In both cases, however, these were rather small gains.

A direct comparison between records made on the two intelligence tests is furnished in Tables III and IV. Each testing was tabulated in securing these figures, that is, if a pupil took the test three times, he was counted three times. He was counted in each

age-group to which he belonged, when he took the National Intelligence Test. It happened in a number of cases that the Binet-Simon and National Intelligence Tests were given several months apart. In all such cases we have multiplied the pupil's chronological age at the time the National Test was given by his intelligence quotient (Binet-Simon) to find his Binet-Simon mental age. Table III shows the median chronological age for the several chronological-age groups. This table also shows the median Binet-Simon and National mental ages for these groups. From these data it is estimated that on the average the National Intelligence Test yields a mental age approximately seven months

TABLE III. NATIONAL MENTAL AGES CORRESPONDING TO BINET-SIMON MENTAL AGES

Median Chronological Age	Frequency	Binet-Simon Mental Age	National Mental Age
8–5	35	9– 3	9– 7
9–6	54	10– 7	11– 0
10–6	64	11– 9	12– 3
11–5	63	12–10	13– 6
12–4	67	13– 8	14– 6
13–7	58	15– 0	15– 8
14–5	46	15–11	16–10
15–3	27	16– 6	17– 2

higher than the Stanford Revision of the Binet-Simon Scale. Table IV shows a somewhat similar condition with respect to intelligence quotients. The first column in Table IV gives the grouping we selected for convenience in tabulating our Binet-Simon intelligence quotients, and the second column shows the median intelligence quotient (Binet-Simon) which fell within each group so determined, while the third column shows the median National intelligence quotient. Thus the median Binet-Simon intelligence quotient for the group ranging from 121-130 is 124, while the corresponding National intelligence quotient is 133.

These pupils have received a great deal of practice in taking tests, and are probably as near practice-proof as children in the grades ever become. Our data, in all likelihood, indicate the

differences which may be expected when children are test-wise. Unfortunately, we have no data to offer as proof that this difference is due to practice in taking tests and not to a defect in standardization. In view of the fact that we found an average gain of 3.7 in intelligence quotient for the National Intelligence

TABLE IV. CORRESPONDENCE BETWEEN THE INTELLIGENCE QUOTIENTS OF BINET-SIMON SCALE AND NATIONAL INTELLIGENCE TEST

Binet-Simon I. Q.	Median Binet-Simon I. Q.	Median National I. Q.	Difference
131+	135	147	+12
121–130	124	133	+ 9
111–120	115	122	+ 7
101–110	106	114	+ 8
91–100	95	101	+ 6
81– 90	84	81	− 3
− 80	78	69	− 9

TABLE V. CORRELATIONS BETWEEN SUCCESSIVE SCORES ON THE NATIONAL INTELLIGENCE TEST AND ON THE BINET-SIMON SCALE
N = 131

	National Intelligence Test			Binet-Simon Test		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Test 1		.90	.91		.88	.91
Test 2	.90		.93	.88		.92
Test 3	.91	.93		.91	.92	

Test during a period of twenty months and of 1.7 for the Binet-Simon Scale, we feel safe in assuming that the difference is due to practice.[2] When pupils' records found in successive tests are correlated with previous records, we find very little difference in the correlation ratios given by the Binet-Simon data and those given by the data of the National Intelligence Test. We do not know, however, how far our correlations given in Table V are

[2] We have Binet-Simon data on 62 cases showing a gain of 3.2 during a period of five years.

influenced by the fact that our subjects had previously taken many tests.

Correlations between the National Test and Binet-Simon Scale and other criteria are presented in Table VI. The National Test gives the highest correlation with the Binet-Simon.

TABLE VI. CORRELATIONS BETWEEN NATIONAL AND BINET-SIMON TESTS AND OTHER CRITERIA
N = 131

	Binet-Simon[a]	National[a]	Teachers' Judgments[b]	Educational Tests[c]
Binet-Simon		.72	.55	.58
National	.72		.76	.79
Teachers' Judgments	.55	.76		.79
Educational Tests	.58	.79	.86	

[a]Average of three testings.
[b]Average of three judgments taken annually.
[c]Average of percentile rank of seventeen different ratings.

The educational-test composite gives the highest correlation with the National Test, but the teachers' judgments give almost as high correlation. Teachers' judgments and educational-test composite give the highest correlation, but the judgments of the teachers were probably influenced by the fact that they were familiar with the scores made by their pupils in mental and educational tests.

Conclusions

1. The results presented in this study are not typical, since the subjects had had an unusual amount of practice in taking tests before these data were secured.

2. For the individuals studied here, the National Intelligence Test distributes the scores over a wider range than the Stanford Revision of the Binet-Simon Scale.

3. The Standard Revision and the National yield equally high correlations.

4. The National gives a higher correlation with teachers' judgments and with educational-test scores than does the Stanford Revision.

5. According to these data, teachers' judgments give a higher correlation with educational-test scores than with either the National or the Stanford Revision.

6. For these children, the National yields higher mental ages on the average than the Stanford Revision.

7. Probably it would be wise to give all pupils a certain amount of practice in taking tests before standardized tests are administered.

THE EDUCATIVE VALUES OF PRACTICE TEACHING

JOHN LOUIS HORN
Mills College

Large and influential sections of the American faculties of liberal arts still adhere to the theory that education consists mainly in transmission and that the educative process consists of learning. Knowledge, they believe, has educational values per se, that only works written in the past are classic, and that only those who thought long ago are masters.

Students of education hold that knowledge which does not function in the lives of the students is not educative, while everything which influences conduct and enriches life does have educational value. They also hold that every aspect of the curriculum and method should be judged in terms of specific objectives and deemed effective only in so far as these objectives are attained. They assert that the purpose of education is the ennoblement and enrichment of life as a direct result of the educative process. But many college teachers have not yet heard of these ideas or, at least, have not admitted them.

These same college and university teachers complain that their students of English enjoy the magazines and the plays adapted to the tastes and intelligence of the uneducated, and are adept at the speech mannerisms of the illiterate. Students, they say, are not distinguished for their taste in music or art and spend their leisure unworthily. These faculty members, however, continue to believe that content will function—a belief which is never subjected to the test of reality. They continue to teach with a view to absorption and accumulation rather than growth and development; and are quite undisturbed by the evident failure of this process to influence conduct.

Against such inertia and resistance to the new—because it is new—Departments of Education in American institutions make difficult headway, particularly when they do not overemphasize

the claim of cultural or disciplinary content, but undertake to function frankly as teacher-training instrumentalities.

The writer recently found it necessary to ask a college faculty that academic credit counting toward the degree be granted for work in practice teaching. Hitherto credit had been granted on the records to be presented to the State authorities in fulfillment of certification requirements, but as this credit was not recognized academically it could not be counted toward the degree, and had to be carried by the student in addition to the degree requirements.

The request for academic credit was based principally on a number of contentions in which an effort was made to show that whatever functioned in an educative manner, influencing conduct, was worthy of credit inasmuch as it contributed to the student's education. Practice teaching is the student's opportunity to test and apply theory and is, therefore, comparable to laboratory or field work in other departments. The development of skills is a recognized factor in education coordinate with knowledge and emotional development, as evidenced by the fact that certain laboratory courses aim merely at technique. The faculty was reminded that certain credit is granted for facility acquired in various arts, and finally, that professional preparation must be frankly recognized as one of the valid objectives of the college.

It was urged that the work resembled very closely certain standard courses for which credit is given without challenge, as for example, work in debating in the Department of English. Practice teaching is, as a matter of fact, almost indistinguishable from this type of academic course. General preparation for a particular debate or subject of instruction, preparing the brief and planning the lesson, delivering the final address and doing the actual teaching are essentially similar activities. To those who objected that practice teaching had no content, by which they meant new knowledge and information, reply was made that neither had the work in debating. Obviously this work aims at the acquisition of skills, and not at addition to the student's stock of information.

The writer went one step further, pointing out that there was one distinction between practice teaching and debating: whereas the student-teacher meets a real situation, teaches real children in a real school, carries real responsibilities, and is influential in real outcomes; the student in the debating course meets only mock situations, and nothing depends on the outcome. To make debating as worthy of credit as practice teaching, it was contended, members of the debating classes should be required to have themselves elected to legislative bodies—municipal, state, national, or otherwise.

In addition to this series of a priori arguments, one objective means of evaluating the educative utility of the work in practice teaching seemed available. The student himself is continually agitating in his mind the educative value of the various phases of the college course, the college environment, the college personnel, and the college method. He is, after all, the focus of all the educative efforts. The professor is frequently more of a student than he is a teacher; he is naturally biased in favor of his own subject; he has rarely if ever studied the other man's subject from the point of view of its educative value; he not infrequently tends to think in traditional terms; and finally, the fact that he turned out a professor indicates that in most cases he never was a typical undergraduate. It seemed to the writer that the college Senior, after four years of the process, might be a highly reliable witness concerning the educative value of any element of the curriculum to which he had been submitted.

It was, therefore, decided to go to the student for evidence regarding the wisdom, propriety, and justice of granting academic credit for the work in practice teaching. Thirty-four students were certificated last spring, in the college under discussion. A request was sent out through the various departments that all students who had done practice teaching as part of their preparation should assemble at a certain time and place. Short notice was given, and no method of compulsion could, in the nature of things, be employed. Thirty students responded. All were Seniors or graduates, and practice teaching

had, in each instance, been taken in either the senior or the graduate year. The following self-explanatory questionnaire was distributed:

This questionnaire represents an effort to determine whether or not practice teaching is a subject that is fully worthy of academic credit counting toward the degree. You will not be required to sign this paper, and you are earnestly requested to place loyalty to your college before loyalty to your department.

If you answer the three questions which follow in an earnest and sincere effort to help, you will probably contribute the most sure and certain information that is available.

1. Should academic credit be granted for practice teaching?
   Yes. No.
2. Give a one-sentence reason for your answer.
   .....
3. If in the preceding questions you have favored granting academic credit for practice teaching, give a series of phrases to indicate what the experience has meant to you in personal educational development and values.
   .....
   .....
   .....

The first of these questions, whether or not academic credit should be granted for practice teaching, was answered affirmatively by all of the thirty students who participated. Presumably they felt that there were sound educational reasons for considering this work on a par with the rest of their college experience.

The returns from a request to give a one-sentence reason for the affirmative opinion that academic credit should be granted for practice teaching are, obviously, difficult to report with mathematical accuracy. Yet it was felt that a type of question which would facilitate tabulation would have the greater defect of suggesting values. Besides, the attempt to list values would in the nature of things have been made out of whole cloth and would have involved a theory or bias regarding practice teaching. The method employed, while making it difficult to tabulate results, is advantageous because it places the entire burden on the person answering the questionnaire.

The results finally condense down to six classes. In a number of instances, obviously, more than one type of answer occurs in a given sentence:

1. Two replies frankly emphasize the practical—professional training—value of the work, the fact that it will be useful after college. "It is by far the most practical and worth-while thing to be gotten out of college," reads one of the replies.

2. The utility of renewing one's knowledge gained in other courses, appealed to four of the students as inherent in practice teaching. It requires one to summarize, employ, coordinate, test, and renew knowledge already acquired. "It helps one," reads one reply, "to a thorough grasp and understanding of *all* the knowledge one has acquired in a college course." Six other replies, reason much along the same line, basing their opinion on the fact that the course gives opportunity for the application of principles of psychology and pedagogy previously learned, —the laboratory aspect of the situation.

3. Thirteen replies take their stand on rather simple grounds of fairness, basing their opinions on the amount of time, effort, and preparation required. Yet it would not be proper to conclude that these persons were disregarding the educative value of the work. "It is as educational," reads a typical answer, "and demands as much intelligent preparation, as any course in the curriculum."

4. Six of the replies content themselves with the comment that the work has educational value similar to other of the courses in the college. "It has been a functioning part of my educational experience."

5. A number of specific values are referred to by individual replies, which it will be interesting to list:

a) Furnishes opportunities for contact with, and understanding of, people as preparation for living.
b) It was the one thing, except similar work in sociology, in the four years, which develops, tests, and makes one always on the alert, brings out initiative, and puts one on her mettle.
c) Gives scope for self-expression, control, teaching ability, responsibility.
d) Involves real constructive thinking.
e) Necessitates the adaptation of theories to suit occasions.
f) "In practice teaching you give,"—as opposed to "accumulating ideas."
g) Gives an insight into American life and institutions.

6. We may, finally, note four particularly enthusiastic replies:

a) The course in practice teaching is "more valuable than any other in its development of initiative and constructive coordination of college work."
b) The work requires "more power, time, and initiative, than any academic course that I have taken."

c) The work is "far more valuable as an educational factor than many of the academic courses offered."
d) It is one of the most valuable courses that I have ever taken.

The answers to the final request, calling for a series of phrases to indicate what the experience had meant in personal, educational, and developmental values, doubtless encouraged a considerable amount of overlapping. So the returns proved, and yet this overlapping has some value in that it tends to emphasize the most important aspects of the work in practice teaching as viewed by the group. The difficulty of tabulating the replies, referred to at the beginning of the last section, applies here as well. These will be condensed under seven headings. It is obvious from the form of the question that any given student's answers may be scattered through all the headings. Wherever possible, the number of times a particular value is mentioned has been indicated.

1. We may begin with the obvious, immediately applicable and utilitarian aspect of the replies, the professional and after-college aspect:
a) Eleven replies emphasize the fact that the work improves one's professional ability.
b) Ten replies emphasize the fact that it gives one a glimpse of one's future profession; one that it broadens one's professional outlook.
c) Three replies emphasize the fact that the work gives one a desire to work and help improve public education; two that it gives one a realization of the teacher's point of view; three that it gives one an appreciation of the seriousness, importance, and obligations of the profession; three that it clarifies the meaning of the great problems of public education.
d) One student said that the work tests one's professional ability, and two, one's knowledge, while three believed that it gives one the opportunity of determining whether one has an aptitude and liking for the profession.

2. The group of replies now to be presented refer to particular qualities which the work develops in the student. The figures indicate the number of times the particular quality is mentioned. Practice teaching then develops:

Self-confidence or reliance	20
Initiative, resourcefulness	9
Poise	7
Self-control	7
Leadership	5
Interest in civic and social problems	5
Sense of responsibility	4

Tactfulness 3
Personality 2
Character 1
Power 1
Ingenuity as opposed to the hitherto needed quality of acquisition 1
Cooperative spirit 1
Sense of fairness 1

3. The work improves one's studies, 1; social development, 2; English, 3; and concentration, 4.

4. This work gives one:
An understanding of childhood and in that sense preparation for motherhood 7
The satisfaction and opportunity of applying knowledge 7
Interesting, worth-while, profitable contacts 6
Knowledge of people, human nature 5
Scope for self-expression, the development of latent talents 3
The sense of rendering service of a substantial, useful character 3
A sense of doing something tangible 1

5. Practice teaching is valuable in that one gains by:
The clarification of subject-matter, ideas, knowledge 4
The chance to test educational and psychological theories 3
The criticisms received 1
The coordination of one's knowledge 1
Growth through service 1
The need for paying attention to one's personal appearance 1

6. The work gives one:
Higher ideals
New experiences. "It has been my most vital experience in college," reads one of these replies 3

7. Finally, the work in practice teaching is valuable because it:
Requires one to review his academic work.
Makes one's work more purposive.
Motivates all the rest of one's work and study.
Supplements academic work with the social, the human element.
Gives one a chance to experiment.
Requires clear and quick thinking.
Helps one to an appreciation of the individual in society.

GRAMMATICAL KNOWLEDGE AND READING ABILITY IN FRENCH

ARTHUR G. BOVÉE
School of Education, University of Chicago

To attempt to reach a definite conclusion as to the exact relation between grammatical knowledge and reading ability in the space allotted to this article would indeed be ambitious, if not rash. If, however, the issue can be raised or a question which many had thought settled can be reopened, this will be an ample reward for the time spent in this study.

The main reason for interest in this question is the fact that no conclusive evidence has been adduced proving that conspicuous success has attended the teaching of French in the United States. It is, therefore, proper to conclude that we are not warranted in viewing with any degree of complacency our present prodigious expenditure of time and effort and the attendant results. Consequently, since the question has a direct bearing on the results we obtain, it is worthy of most careful consideration, especially in view of the almost hostile attitude of certain educational specialists.

Before proceeding to the main discussion of the subject, let us consider for a moment what a language really is and what our aims are in teaching it. To state that the fundamental purpose of language is the expression of human thought and that the natural function of grammar is to aid the more perfect expression of the thought is axiomatic. The flesh and blood of language are words combined into groups expressing complete ideas. Obviously thought content is the principal element in language; the question of form is a secondary consideration. Grammar cannot be considered as an end in itself, but merely as a tool.

With this definition before us, it is pertinent to inquire what is the objective to be attained in the teaching of the French language. Professor Purin has said:

> The chief aim of a high-school course in a modern foreign language should be the acquisition of a ready and accurate reading knowledge.

. . . . By reading knowledge I understand (1) the ability to pronounce correctly, fluently, and with proper intonation; and (2) the ability to understand, without recourse to translation into the vernacular or to the dictionary, the content of the piece read.[1]

This clear statement of the objective not only represents the consensus of opinion of most teachers, but also is the only one which can possibly be defended, with any hope of success, against the attacks of educational specialists. From the foregoing we may conclude that the main objective is the power to get the thought directly from the printed page, and that the main service we ask from the grammar is to help in the better attainment of our objective.

If our statement of the objectives is valid, we must evidently be guided by it, not only in the choice of grammatical material, but also in the method to be employed in teaching it. We may properly inquire at this point if the ability to reproduce in order the principal parts of an irregular verb is an absolute *sine qua non* for reading. Can the pupil read with understanding and yet be ignorant of the rule that the definite article is omitted from the partitive construction when the noun is preceded by an adjective? Just what grammatical material is the most effective equipment for reading purposes has yet to be determined. I entertain grave doubts whether the pupil is aided to a more perfect appreciation of the thought value of French words if he is taught to consider language as a series of examples of laws of grammar. Do the terms partitive, disjunctive, conjunctive, past indefinite, and pleonastic, illuminate his mind? These and many others are the questions I should like to raise, for upon the answers depends the whole question of the place of grammar in teaching, as well as the method of teaching it. May I suggest that observation of the way in which a child learns to read in his own language as well as of the processes of pupils in the acquisition of reading ability in the foreign language be used as guides to investigation. These are just a few of the questions to be settled before reaching a final conclusion on this matter.

[1] Purin, C. M. "Recommendation for a Standard Minimum in a Four-Year High-School Course in French, German, and Spanish," *Modern Language Journal*, 6:297 ff, March, 1922.

With a feeling of considerable caution I approach the main point of this discussion, because it is fertile in its possibilities of being misunderstood. I should feel that my efforts had been more than futile if any one should receive the impression that I do not believe in the study of grammar. The problem is not a question of grammar or no grammar. It is a question of the most efficient way of studying grammar in order to promote the acquisition of a good reading knowledge of French and thus to attain more speedily and effectively the ultimate objective of the study. Herein lies the crux of the whole matter.

We derive our opinions from our experience; as time goes on these opinions ripen into convictions. My opinion which is approaching the point of a conviction is that the grammatical knowledge acquired through the practice of translating English sentences into French does not increase one's power to get the thought from the printed page. This method reposes on the unproved assumption that the more of this type of work we do, the better the pupils read. The problem before us then, is to try to determine to what extent grammatical knowledge acquired as heretofore described becomes an active factor in developing the ability to recognize thought. With this problem in mind, tests were devised which would enable us to investigate the relation between grammatical knowledge and reading ability. In order to have a better understanding of the data which is to be exhibited a brief description of the tests will be necessary.

In the first place, a test was constructed consisting of a series of forty questions, given orally, ten for each year of a four-year course, so formulated that a one-word answer either in English or in French would give unmistakable evidence of comprehension. These forty questions were arranged in order of difficulty, the order being determined first by guess and subsequently rearranged in accordance with the results obtained from several hundred trials. The purpose of this arrangement was to obtain a distribution with its obvious values. Here was a basis for the selection of the questions. The first ten consisted of concrete vocabulary, based on the usual facts of everyday life; the second ten were

based on the usual idioms of the French language; and the third ten were constructed with the vocabulary usually found in French short stories of Daudet, Maupassant, Coppée, such as are generally read in the third year. The fourth and last set of ten questions were based on the well-grounded assumption that the pupils of the fourth year should possess some cultural knowledge of France, politically, historically, geographically, and socially. When this first test had been perfected by sufficient experiment and the order of difficulty well established, the vocabulary of the forty questions was woven into a connected narrative which, for the purposes of scoring, was divided into eighty thought units. This, then, is the history of the development of the test which in its second form was handed to the pupils as a visual translation test whose objective was to explore the ability of the pupil to get the thought and tell us at what point he reached his limit. This is called Test E.

In order to construct the applied grammar test a canvass was made of the principles of grammar which seemed to be essential to a good reading knowledge because of their frequent recurrence. The following represents the divisions of the subject which were chosen: the article, agreement (adjective, participle), pronouns (all types), verbs (tense formation, personal endings, tense usage, moods), prepositions, and negations. The material was arranged under these headings in the form of completion or substitution exercises, translation of one word with the rest of the sentence given, or changes in the person or tense of a verb. It was impossible to cover all the grammar a pupil can apply, but the choice is quite representative. The fact, that the knowledge explored by this test is usually acquired by two years of high-school work, might well be added. This is called Test C.

These two tests, E and C, were given in three Chicago high schools which we shall designate as School G, School D, and University High School. In considering the results of the tests, we are not at all concerned with the scores made, but only with the relation of growth in reading ability, or rather translation

ability, as measured by Test E, and with the growth in grammatical ability as measured by Test C.

School G on the charts is a large city high school using the grammar-translation method and offering a four-year course. The teacher in charge of the advanced courses enjoys an enviable reputation as a technician.

School D is another large city high school using that form of the direct method which emphasizes conversation. The third school is the University High School which uses the direct and phonetic method.

Figure 1 offers a graphic comparison of the results of the tests in School G. The striking thing shown by the curves is the fact that during the fourth year there is a marked increase of grammatical knowledge, but that this improvement is not accompanied by a corresponding increase in the power to grasp the thought.

Figure 2 shows a similar comparison for school D. The lack of correlation between improvement in grammatical knowledge and improvement in ability to read the thought is as marked as in school G, but the relation is the inverse of that found in the latter. G, which uses a grammatical method, shows consistent improvement in grammar with a marked upshoot in the last year, while its improvement in reading is at a standstill during the fourth year. School D, in which the direct method is in use, shows a marked falling off for grammar in the last year, the third, while its reading curve has the upshoot in the last year. Further it scores 5 points higher on median achievement at the end of the third year than does G at the end of the fourth.

Figure 3 exhibits a comparison of the results of the grammar test in the University High School, which employs the direct method, with those of school G using the grammar method. The obvious superiority of the University High School may be partly explained by more favorable teaching conditions. Let us dismiss the issue by conceding the point at once. But what cannot be so easily dismissed is the fact that a group of over three hundred pupils have acquired a considerable body of grammatical knowledge without being taught grammar as grammar. To them

the French language was a vehicle for the expression of thought. Their grammatical knowledge came mainly as a result of their observations of the variations in forms for the purpose of expressing different thoughts. It is interesting also to contrast the

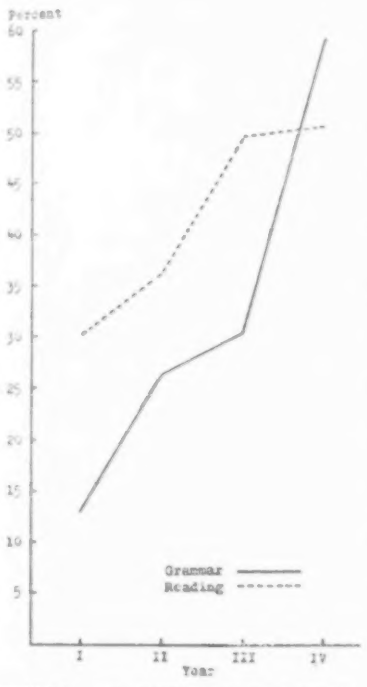

FIGURE 1. RELATION OF READING POWER TO GRAMMATICAL KNOWLEDGE IN FOUR GRADES OF SCHOOL G

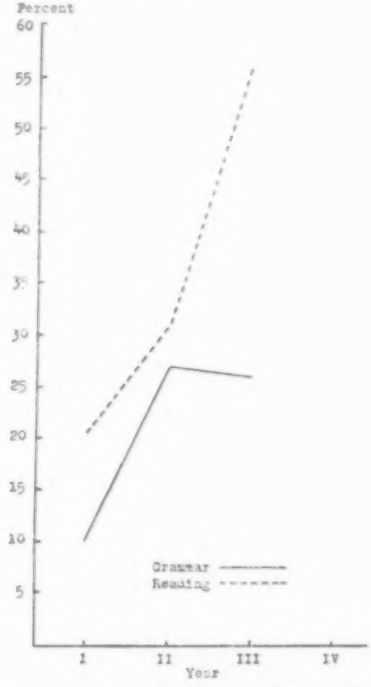

FIGURE 2. RELATION OF READING POWER TO GRAMMATICAL KNOWLEDGE IN THREE GRADES OF SCHOOL D

character of the curve. The curve of University High School shows steady and consistent progress. The curve of school G approaches the horizontal decidedly during the third year, but indicates a decided increase of knowledge during the fourth year.

Figure 4 gives us a picture of the results of the three schools in translations. The outstanding fact here is not only the superiority of the two direct-method schools whose curves have

a similar tendency, but the marked horizontal character of the curve for the grammar-method school.

It would be presumptuous to make any extensive claims on such meager data. These results in themselves may be merely highly suggestive of important possibilities which merit investiga-

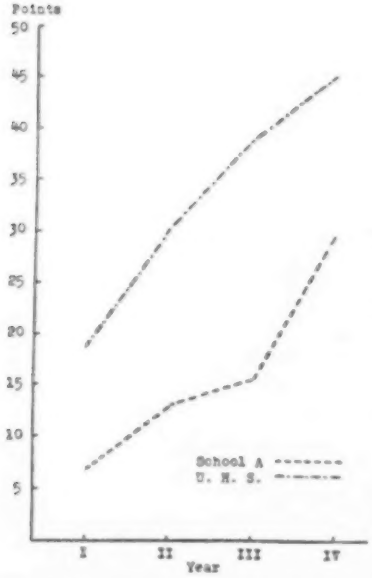

FIGURE 3. RELATIVE ATTAINMENT IN GRAMMATICAL KNOWLEDGE IN TWO SCHOOLS EMPLOYING DIFFERENT METHODS

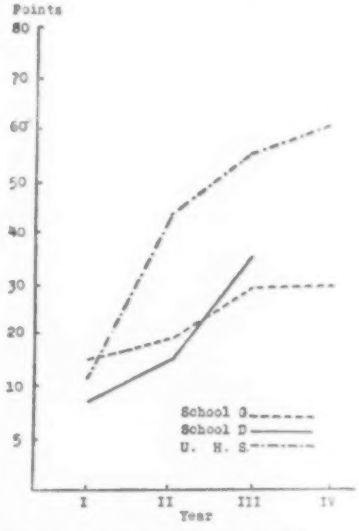

FIGURE 4. RELATIVE ATTAINMENT IN READING IN TWO SCHOOLS EMPLOYING THE DIRECT METHOD AND IN ONE SCHOOL USING GRAMMATICAL METHOD

tion. I repeat, as an isolated affair the value of these results may be doubtful, but when considered along with a similar and more extensive study of ability in Latin in secondary schools, conducted by H. A. Brown in the schools of New Hampshire, they immediately become highly significant, since the conclusions are identical. In this report we read the following passages:

> In other words, pupils have no greater mastery of grammar in those schools which teach the grammatical method in the first year in its extreme form, which have a systematic study of grammar in connection with weekly exercises in prose composition, and which give daily atten-

tion to syntax, with constant use of a textbook on grammar, along with the study of Latin authors. In view of this fact, the large amount of the pupils' time and the teachers' effort given to the direct study of grammar is unjustifiable. Here is an important source of waste in teaching.

Our observation of the work of these eight schools convinces us that results in ability to get the thought of the Latin, which are so meager, would be less so if a shift of emphasis should be made from grammar to translation or reading of Latin. In a good many schools, half the time of the recitation period is often spent in discussing the minutiae of form and syntax, which according to the evidence contributes almost nothing to either ability to translate or read, or knowledge of construction. It is moderate language, then, to say that the waste here is extravagant. Public funds are literally being thrown away in the expenditure of this large amount of time and effort which results in nothing of value.[2]

The fact that a school stands high in construction does not necessarily insure a high standing in ability to get the thought from Latin. Schools with high and low scores in grammar are found in all divisions.[3]

The final conclusion of this whole problem is that a high score in knowledge of construction does not necessarily insure a corresponding great ability to apprehend the thought of the Latin rapidly and accurately, although there is some correspondence, in general, with exceptions in particular cases.[4]

Later we find the following:

What, now, shall we say concerning the practice of spending so large an amount of time for three years on prose composition for the purpose of clarifying and fixing principles of grammatical construction? What shall we say concerning the tendency in one-third of the schools to spend a considerable amount of time and in another third to devote a very large amount of time to direct teaching of grammar? The facts serve only to strengthen the conviction stated in an earlier chapter that the excessive amount of time and attention given to grammar is an unwise and wasteful practice.[5]

Finally, President Brown has this to say:

We have also further proof of the futility and wastefulness of that plan of teaching which spends so large an amount of time learning paradigms and rules and gives so much attention to syntactical analysis of the Latin texts which are translated in class.

One of the interesting and rather significant facts in connection with this aspect of the investigation is the fact that the schools of

[2] Brown, H. A. *Latin in Secondary Schools.* Oshkosh, Wisconsin, State Normal School, 1919. p. 91 ff.

[3] *Ibid.* p. 93.

[4] *Ibid.* p. 94.

[5] *Ibid.* p. 97.

Group I which devote least time to the study of Latin grammar in a formal way, and, in fact, least time to it in any way, with a conspicuously smaller time allotment in general, get on the whole as satisfactory results in Latin grammar. This is especially significant and is an indication of the source of the great waste which takes place in the teaching of Latin.[6]

The final sentence raises the real issue. It is the question of waste. Educational specialists are busy pruning the curriculum of all dead wood. Branches of study which are decaying through the process of formalistic rotting will be cut off ruthlessly. If the teachers of French cannot give a good account of their stewardship, there will be a day of reckoning. If the study of French is sufficiently valuable to warrant a permanent place in American schools, it behooves the teachers of the subject to be sure that the time allotted to them is spent in the most economical and efficient way, for the day is not far hence when they will be on the defensive. Complacency regarding this matter will be sure to entail the most serious consequences.

[6] Brown, *op. cit.* p. 115.

Editorials

AN OPPORTUNITY FOR SUPERVISION

Have you ever noticed how lazy you are when you have a chance to be? How you observed your proneness to take advantage of an easy day by doing poorly the things you would do well on one of your high-pressure days? Haven't you found, too, that if you want to get a thing done—an act of citizenship performed, a letter written, or an interview granted—you have to go to a busy man to get the thing accomplished?

Then there is the working man. He's an object of our concern. We argue for an eight-hour day for him and then for a six-hour day. And what is the basis of the argument? Is it not that he may make better use of the time which he will thus have at his disposal—better use for himself and his family? Is it not that he may get more sweetness and light, that he may be a better man? Sometimes he does, indeed, make better use of his leisure. Some workers do undoubtedly become better men through improved working conditions; very likely enough of them do to justify the argument. But too often the worker does a poorer job than he did before; and in his leisure time instead of getting sweetness and light he gets drunk.

School people have somewhat the same disease—Oh no, not inebriety, but malingering. There is no use protesting, no use whatever in quarreling with a very human tendency. Work is distasteful until one gets into the swing of it and there isn't any swing unless there is a certain pressure, a certain "go."

We think this has been left out of account in many an argument. Consider the case for small classes. Upon what does it rest? Doesn't it rest essentially upon the idea quite generally expressed by superintendents, principals, and teachers: "Give us

small classes so that we may pay attention to individual pupils, so that we may make a better adjustment of the school to each particular child." But do the teachers of small classes avail themselves of the opportunity to do the things they imply they will do if they are given but few pupils to teach? There is a good bit of evidence that they do not.

The theory that they can make better individual adjustments is all right, but the ordinary demands of school teaching do not actually exact such adjustments. The superintendent who has succeeded in securing smaller classes does not *have* to demonstrate that pupils are better taught. He can content himself with sweet reasonableness. He can assert the greater opportunities that his teachers now have; and along this line he can give many convincing details. Similarly the teacher finds it unnecessary to be specific in proving her work to be superior. The standards and traditions of craftsmanship in teaching permit her to "get by" without such proofs. Frequently, therefore, the only visible result of reducing the size of classes is that the teaching job becomes less exacting. As to work with individual pupils, the teacher may—and often does—do less of it than before.

Consider also, the theoretical argument for the grouping of pupils according to ability. We say we can make better adjustments to individual needs when these needs are not greatly different. The theory is almost unassailable. Better teaching should result. The evidence, however, that it does result is by no means clear. Homogeneous groups have been known to do no better than heterogeneous groups. The *opportunity* is better, of course, but the opportunity may not be seized.

As in the case of the small class or the homogeneous group, the situation may be handled in either of two ways, with gradations between them. Either the teacher may do less work and produce the same result as before, or she may do the same amount of work and obtain a superior result. The *argument* for small classes and for homogeneous grouping assumes the latter outcome, namely the superior result. The evidence, however, suggests the former outcome, namely the same result as before.

Indeed, one may reasonably expect this to be the case unless precautions are taken to prevent it.

These precautions are of a supervisory character. Supervision must intervene to secure the advantages which superior organization and larger outlays of money may have brought about. And this supervision may be supplied by the teachers themselves. If the traditional type handed down from above is thought to be inferior, let the teachers take themselves in hand and assure to the children and to the public the outcomes which should result from greater public effort. Why should not teachers unite for improvement as well as for advancement, why not for artistry as well as salary, why not for expertness as well as tenure?

At any rate, whether supervision is of the democratic or of the aristocratic kind something additional to better conditions of work must be provided. And when research workers go about it to measure the effect of small classes, or ability grouping, or segregation let them take supervision into account. Let them not, for example, assume on the basis of tests of performance that small classes are of no advantage. If the test results show no advantage let them be content to say that and no more. If they must condemn, let them condemn, the lost opportunity.

If small classes are organized, the additional step must be taken of seeing that the better sort of teaching which small classes presumably permit is the kind of teaching actually done. If homogeneous grouping is put into operation, supervision, whether of the self-starting variety or otherwise, must be invoked to see that the fruits of this supposedly superior class organization are secured. The teacher of the gifted group must know how to teach gifted children. Not many teachers do know how. The teacher of less gifted groups must make similar adjustments. As things now stand, the teacher who succeeds best is the one who with the usual range of pupil ability at hand adjusts her teaching to a level somewhat below the average. This is the level to which most curriculums and most books on method are adjusted. Moreover, it is the level to which the teacher as the years of experi-

ence go by becomes accustomed. What right have we to expect her to adjust herself to other levels and to a narrow range of ability simply because our theory tells us that better teaching ought to result under such conditions?

We discern in this a more than usually valuable distinction between administration and supervision. Administration may supply the improved working conditions, but it is supervision which must secure the results of the administrative measures.

B. R. B.

Once upon a time a very good friend of ours criticized the Review Department of the Journal because some—or did he say "most"?—of the books were old by the time we reviewed them. By way of drawing the deadly parallel he mentioned another periodical whose reviews were issued so promptly that after reading them one could talk knowingly about the latest books and make quite an impression.

Now what we are trying to do is to publish good reviews rather than quick reviews—able expressions of opinion from people whose opinion is worth having. We should like to have the person who *must* make a show of his reading brag about the review as well as the book.

Of course, we shall prefer to bring out timely reviews; but as between something hastily made up by the editor or a graduate student or a printer's devil and something written by an expert who takes time to read the book, the choice is as plain as the choice between a twenty-dollar gold piece and a plugged nickel.

We have a notion, too, that a review should serve as a guide to reading. Now, in the wild welter of books, bulletins, reports, brochures, monographs, pamphlets, and reprints, to say nothing —positively nothing—about charts, job-analysis sheets, budgets, schedules, and such ephemera—amid such surroundings how shall a man pick his way? Who wants to read the latest books? At least, who would not prefer to read the best books? And the best books are never evident at first. Everybody knows the rôle of

time as a critic; and some of us who are a year behind in our reading have been saved from much rubbish.

All this is intended to give a proper setting to three reviews of 1923 books which we unblushingly publish in this issue. There's one by Doll on Kohs' *Intelligence Measurement*. This is a keen analysis, cleverly written. Its high point occurs where the reviewer uses the author's method to prove that his Block-Design Test is a vocabulary test! And then there's one by Goddard. You will really want to know what he has to say about Downey's *Will-Temperament and Its Testing*. His *obiter dictum* on those who denounce mental tests as compensation for their disappointment in not having originated them is refreshing. His invocation of the spirit of Binet—and who knows that spirit better than he?—is the kind of thing you don't find in a review very often. Finally—since *aller guten Dinge sind drei*—there's a review by Swift (F. H.) of Smith's *The Bureau of Education: Its History, Activities, and Organization*. Professor Swift is just the man to give this monograph a proper setting. He does so admirably. You may think, before you read this review, that we have let him have too much space; but you won't think so after you read it.

So that's why we print at considerable length reviews of three 1923 books. We do so for the only reason that justifies us in printing anything—namely, because we think they are interesting.

Does some psychoanalyst mutter "defense reaction"?

B. R. B.

Reviews and Abstracts

S. C. Kohs. *Intelligence Measurement.* New York, Macmillan Company, 1923. 312 pp.

This volume is in two parts, one on the theory of intelligence and its measurement, the other a standardization of a new mental test, the Kohs Block-Design Test. These two aims are not so closely correlated as might be, since the theoretical argument is not, as presented, essential to the test standardization, nor is the standardization essential to a demonstration of the theory. Moreover, a great deal of more or less irrelevant argument tends to distract attention from both these issues.

The main burden of the book is devoted to the Block-Design Test, its presentation, standardization, and justification. The author's meticulous care in describing his subjects and test materials, definitizing his procedures, elaborating his statistical treatment of data, and validating his results, we commend most heartily to past and future standardizers of mental tests. If the treatment at times seems unnecessarily or confusingly minute and detailed, it is a good fault, that might well be committed by others in this field.

The test is of the "performance" variety and requires the subject to reproduce designs in color by use of colored cubes. The results are found to correlate with other criteria and measures of intelligence in satisfactory degree. The standardization is found to meet the several statistical criteria adopted by the author. The mental processes involved are found to be in accord with psychological theory. All this to the author's satisfaction but not necessarily to the reader's. Thus, although the author recognizes that the normal curve of distribution of test data is no guaranty of test validity, he employs this criterion. Although he decries the limitations of the Binet Scale and the language- and group-test measures of intelligence, he uses these as important bases in validating his test. Although he denies that a diagnosis of feeble-mindedness may be made by mental-age tests alone, he concludes that his test is frequently a differentiating test on border-zone cases. He even goes so far as to state that because two different measures correlate highly they, therefore, measure the same thing! Thus, if the Binet measures intelligence so does the Block-Design Test, because they correlate to the extent $r = .80$ (*circa*). But the test also correlates $r = .77$ with Terman vocabulary. Is it, *ergo*, a vocabulary or a language test?

The principal objection one might make against this test itself, while tentatively welcoming it into the family of much-needed non-language tests is that the author completely ignores the element of practise. And yet experience with "performance" tests and with this test in particular (which is a house-

hold occupation of very young children in many families, the material being on sale in department-store toy departments) leads to the suspicion that familiarity with the blocks and design reproduction would, at least among normal subjects, greatly facilitate reaction. Moreover, in spite of the unusually elaborate treatment and analysis of results and the exceptional care in experimental procedure, there is no mention of reliability coefficients obtained from re-testing, nor is the validity of the test examined from the point of view of internal correlation.

The theoretical sections of the book are not systematically or logically related. The separate arguments are in themselves valuable so that the lack of coordination is doubly unfortunate. Moreover, much of the argument, while important in itself, is irrelevant to the main issues. Nevertheless, it is most encouraging to see this comparatively isolated attempt to theoretically justify a mental test. Would that others in this field might emulate this example. In this discussion the author makes several original contributions of importance which we hope will not be lost because of their specific setting. Thus the discussion of the percentile standardization of tests, while disconcerting to those who have used its most simple form of application, should not pass unseen or unsung. In the same way the candid and conscientious examination of test validity in terms of numerous criteria sets an excellent example for others to follow. Here again, however, one notes inconsistencies. Thus the author is particularly naive in assuming that performance in his test proceeds by way of the analytic-synthetic process which is his criterion of intelligent behavior.

Of particular interest to the present reviewer is the author's argument (in twenty counts) dealing with the question of the average mental age of adults. But because our own view of this moot question is under fire we feel constrained to pass judgment in this matter to less prejudiced readers. It seems paradoxical, however, that in a volume of such excellent experimental technique and careful analysis the author could also take such a confessedly "stand-pat" position, and rest his case on authority, specious arguments, and speculation rather than on experimental evidence. It speaks ill for applied psychology when scientific conservatism, commendable enough in itself, waits on convenience of professional readjustment rather than on experimental evidence. Had the author been less prejudiced or more candid he must have avoided the confusion into which he is led by a too-critical examination of new data and a too-complacent acceptance of previous theory.

The volume is supplied with an exceptionally good index. Bibliographic references are scattered throughout the volume in footnotes except those used in Part A which are collected at the end of the chapter. An author-index helps to overcome this shortcoming. The use of references is specific to the particular argument and not general to the major topic as contained in the title. The reader who expects a comprehensive treatment of intelligence measurement as such will be disappointed. So also will those who look for an adequate treatment of the whole question of performance tests as a department of measurement specially added to in this contribution.

EDGAR A. DOLL

Ohio State University

DOWNEY, JUNE E. *The Will-Temperament and Its Testing.* Yonkers, World Book Company, 1923. 333 pp.

Ever since the advent of the Binet-Simon Scale for the measurement of intelligence, there has been a more or less constant cry that intelligence is not everything. This comment has had different kinds of significance from the constructive critic who says in effect that measuring intelligence is a valuable thing, but that it would also be valuable if we could measure the emotions, the will, and temperament, down to those who have said that the measurement of intelligence is of no value, that we know nothing about a man until we know his feelings and will.

These different types of criticism of course imply different types of critics. There are psychologists who are especially interested in the affective side of life and consequently seek a method of testing the motions. Then we have the psychologists who might be described by the phrase that used to be so common in biographies, "poor but honest." These men still think that emotions and will can be measured and evaluated without regard to anything objective. Then we have those whose denunciations of intelligence tests illustrate the psychoanalytic doctrine of rationalization and are only crude compensation for their own disappointment that they themselves did not originate mental tests. Had Binet lived to hear these criticisms, both kindly and otherwise, he would surely have met them and his reply would have been one more contribution to modern psychology. Binet is no longer here to write out his thoughts in *L'Anée Psychologique*, but we have only to call up his spirit to learn what he thinks on this question.

To those who feel it necessary to reiterate that the intelligence is not all of the man, Binet's spirit replies: "Certainly not. That is so obvious that I have never thought it necessary to mention it." So far as I know no writer on the subject has ever claimed that in measuring the intelligence he was learning all about the individual.

To the objector who maintains that the measuring of intelligence is not of much value unless we can also measure the emotions and the will, Binet replies in spirit: "Are you not in error in thinking that, because the measuring of intelligence is not everything, it is therefore of little importance? And are you not in still greater error when you are disposed to neglect the intelligence level? Does emotion or will ever compensate for lack of intelligence? Has anyone ever seen a moron with any kind of happy combination of emotion, temperament, and will, such that it entirely neutralized his lack of intelligence? On the other hand the presence of high intelligence enables an individual to control his motions, to manifest this thing that we call will-power, and probably to a certain extent at least, to modify his temperament. May you have success in measuring the emotion, temperament, and will." We need to know the emotions of the intelligent man in order to understand how great is his problem of self-control. We need to know the measure of the emotions of the moron because, since there is little intelligence, the character of his emotions is largely the determiner of his conduct. Heaven speed the day when we shall be able to measure accurately these processes and qualities; but let us not belittle the value of intelligence measurement, for without any question

it is the pilot, the guide, the captain, the chief of mental processes. More than any other process it distinguishes man from animals, the genius from the moron, the efficient from the inefficient. In the words of David Star Jordon "In human life there is no substitute for intelligence."

Dr. Downey's book is an honest attempt to do constructive work. The author seeks to supplement intelligence tests "by other tests for the measureing of factors other than intelligence in the human personality."

The first four chapters beginning with "Human Individuality" and ending with "The Central Idea of the Will-Temperament" constitute a very clear, interesting, and convincing introduction to the book. In the fourth chapter, we are told about the tests themselves. There are three groups of four tests. Of the twelve tests, all but three represent writing reactions. The author's well-known interest in handwriting at once makes it clear why she finds her tests in this field. The Will-Temperament Tests were primarily for individual examining but have been adapted for group work as well, though this requires the omission of some of the tests.

Three chapters are devoted to a discussion of the tests without, however, telling us how the test is given or how it is scored. For this, one must secure the manual of instructions and the record sheets. Some intimation of some of the tests is, however, given in the book, so that one gets an idea of the general procedure. For example, in the test for motor inhibition this power was measured by the length of time the writing of the phrase "United States of America" could be voluntarily retarded. The author states that individual variation in this test is extreme, ranging from a retardation of a few seconds to a retardation of three-quarters of an hour. This test is a great strain on the nerves. The author states that "one woman who maintained her control long enough to achieve a ten score finished on the verge of hysterics, while another was badly nauseated."

The book is well arranged and well put together. The author is herself critical of her own points of view and one lays the book down with the feeling that here at last is a test of the will and temperament which ought to work and apparently does work. Still one has some misgivings. Perhaps the one most serious difficulty arises in the scoring. Where so much depends upon the scorer—in other words, where there are few objective criteria—one feels that probably Dr. Downey can use the scale with valuable results, but that anyone else would probably feel very uncertain of his conclusions. This is, however, no reason for our discarding the test. The subjective element is in everything. If there is more of it in the Downey Will-Temperament Test it only means that those of us who wish to measure these traits will have to work the harder to master the technique. If Dr. Downey can get results with it—and apparently she does—the rest of us can get results to the degree that we approach her in intelligence, and to the extent that we do not let our emotions run away with us. At least, it is a splendid attempt in this most difficult of all fields of psychological testing.

H. H. GODDARD

Ohio State University

SMITH, DARRELL HEVENOR. *The Bureau of Education, Its History, Activities, and Organization.* Baltimore, Maryland, The Johns Hopkins Press, 1923. 157 pp. (Institute for Government Research, Washington, D. C., Service Monographs of the United States Government, No. 14.)

With a view to laying the basis for a comprehensive study of the organization and operations of the National Government, the Institute for Government Research has undertaken the preparation and publication of a series of monographs, of which the present study is one. It may be well to state that the Institute for Government Research is an association of citizens cooperating with public officials in the scientific study of government for the purpose of promoting efficiency and economy in governmental operations and of advancing the science of administration. The Institute aims to bring into existence such information and materials as will aid in the formation of public opinion and will assist officials in their efforts to place public administration upon a more efficient basis.

With this aim, the Institute has in the process of publication a series of monographs giving a detailed description of each of the fifty or more distinct services of our National Government. These monographs are all prepared according to a uniform plan. They present: (1) the history of the establishment and development of the service; (2) its functions described not in general terms, but by detailing its specific activities; (3) the character of its plant; (4) a compilation of, or reference to, the laws and regulations governing its operations; (5) financial statistics showing its appropriations, expenditures, and other data for a period of years; and (6) a full bibliography bearing on the service and its operations.

In keeping with the outline just given, the author of the present account of the Bureau of Education has divided his treatment into four major divisions, as follows: (1) the history of the Bureau; (2) an account of its activities; (3) its organization, and (4) appendices, including (a) an outline of organization, (b) classification of activities, (c) publications, (d) laws, (e) financial statistics, and (f) a bibliography.

The historical account of the Bureau first takes up, act by act, all the legislation affecting the Bureau, from the passage of the Organic Act March 1, 1867, by which a Federal Department of Education was created, down to the Maternity Act of 1921 which provided that the United States Commissioner should be a member of the board established to administer the Maternity Act.

Under the caption "Leadership" an account is given of the work of the various commissioners beginning with Henry Barnard and closing with a brief reference to John J. Tigert, who had scarcely more than entered upon his duties as commissioner at the time the present monograph was written. Dr. Barnard's inability to gain the support of Congress and the other reasons for his failure to increase the influence and prestige of the Bureau are set forth. Of all the commissioners whose work is described, that of General Eaton and that of William T. Harris stand out as most far-reaching and perhaps most significant. An account is also given of the activities of the Bureau and a history of its administrative duties.

Part II devoted to an account of the activities of the Bureau, necessarily duplicates somewhat those sections in Part I devoted to the history of the Bureau's activities. Emphasis is laid upon the fact that throughout its history, the main function of the Bureau has been the collection of data and the dissemination of information. The handicaps under which the Bureau has been obliged to labor are as evident as they will be surprising to one not familiar with its history.

It is unnecessary to attempt to summarize here the description of the organization of the Bureau, or the material contained in the appendices already referred to. The Institute of Government Research and the author of the present monograph are to be congratulated. The account is given in an unusually clear, straightforward style; the facts are pertinent and vital; and the arrangement is logical and helpful. The author has rendered a service of enduring value, not only to students of government and education and to statesmen, but also to that increasing body of citizens who wish to be informed regarding the nature and functions of our agencies of government in order that they may participate intelligently in the duties of citizenship.

FLETCHER HARPER SWIFT

College of Education, University of Minnesota

CUNNINGHAM, BESS V. *The Prognostic Value of a Primary Group Test.* New York, Teachers College, Columbia University, 1923. 54 pp. (Teachers College Contribution to Education, No. 139.)

About two-thirds of this monograph is concerned with the value of the test used—the Printner-Cunningham Primary Mental Test—as an instrument for the prediction of work in school. One study was carried out with forty-three cases in the first grade, the length of time involved being a little over a year; the second investigation, with fifty-four cases in the same grade, for the period of only five months. The results, in both instances, are about what would be expected. Of the first group eighteen showed school work that did not agree with their original test scores, while there were fourteen such cases in the second group. If these figures are compared to the number of children who would be correctly placed if grouped into sections by chance, it is found that the tests were, in the first instance, 24 percent better than chance placement, and in the second instance, 41 percent better.

A slight investigation was made into the reliability of the mental test used. It appears that one application gives a reasonably good indication of standing, that the brighter children vary more than the dull and that these same children profit the most by practice. Since only thirty-six cases were used for this part of the work, the results are merely suggestive.

Throughout the monograph the author has made reference to the over-age cases among the children, of whom there were nine in one group and ten in the other. In all instances she finds that her test does not adequately measure over-age children. This result is in keeping with those of other workers with primary tests; the non-language test does not measure the intelligence of children who are about eight years of age; these children always score much too high.

The monograph contains thirty-six tables of various sorts, and presents the results of a great deal of statistical handling. The reader is somehow left with the idea that, if one applies enough statistics, one may make up for having too few cases to be of significance. Such elaborate handling of forty-three or even fifty-four cases is rather absurd, and the results presented cannot be considered as more than slightly indicative of the trends they are supposed to show.

LUELLA C. PRESSEY

Ohio State University

STRAYER, G. D., and ENGELHARDT, N. L. *Standards of High-School Buildings.* New York, Bureau of Publications, Teachers College. 1924. 95 pp.

This bulletin sets up standards for large city high schools and should be used to supplement the authors' score card for high-school buildings. The score card may be ultilized in analyzing the conditions of a high-school plant, thereby affording a basis for the developing of a building program and a means of checking architects' plans.

In addition to standards for high-school buildings the bulletin also contains many details of interest to the city superintendent including: a bibliography of literature on high-school standards; diagrams and descriptions of athletic fields, shower units, and laboratory equipment; and certain data on the organization and administration of high schools.

The writers' opinion is that the standards can be applied to the best advantage to very large city high schools, but with certain modifications smaller plants can be measured and checked.

P. R. STEVENSON

Ohio State University

ALMACK, J. C. *Education for Citizenship.* Boston, Houghton Mifflin Company, 1924. 287 pp.

The "incidental training for citizenship gained through unguided participation in the life of our democratic form of government" has failed to meet the needs of "the increasing complexity of society." The editorial introduction continues with the statement that supplementing "such incidental training with a study of the forms of the machinery of government" has likewise proved to be woefully inadequate to meet the new needs of society. Real education for citizenship is a long process and involves much of both direct and indirect instruction. The school, in order that it may do its work efficiently, must become "a vast laboratory for citizenship training."

The author, assuming "that good citizenship is the final goal of all teaching and of all school organization," has developed at Stanford University a practical course in Education for Citizenship "to show how abundant are both the direct and indirect opportunities in the modern school and the modern community for civic training and civic experience." He has there developed and emphasized "a method of teaching which has been found productive of good results wherever it has been given a fair trial." This abundant concrete material of the school and the community, together with

the methods of procedure, have been collected and presented in the form of a well-written textbook of 287 pages, which will doubtless prove to "be of great service, not only in teacher-training institutions, but to thousands of teachers now in service . . . . who have not had the benefit of such instruction" as that given by the author in Stanford University.

GEORGE E. CARROTHERS

Ohio University

PETERS, CHARLES C. *Foundations of Educational Sociology.* New York, Macmillan Company, 1924. 447 pp.

It has taken three quarters of a century of development to bring educational psychology to its present stage of objectivity and exactness. Two sets of workers were responsible for this progress, practical educators, interested in the study of children and their educational needs, and scientific psychologists, concerned with the general problems of psychological research. The educational psychologist is a compound of the two and is a comparatively recent product. In the same way educational sociology is a hybrid study appealing to two types of students, the practical educator, interested in the social problems connected with child life as they affect education, and the general sociologist, interested in the school as a social institution. The educational sociologist is a compound of the two and is in the making. That he has a definite contribution to make is evident in this volume by Professor Peters. It is the fourth book to be offered as a text since the writer's "Introduction," published only seven years ago, and is, in many respects, an advance over all previous efforts.

By training and preference Professor Peters represents the educational rather than the sociological approach, and hence his education is sounder and more illuminating than his sociology. He has difficulty in keeping away from the individual as the focus of educational thought whereas the educational sociologist must approach every problem from the standpoint of the social group. His interest must center in the individual only in his collective aspects—his citizenship, his vocation, his religious and cultural activities, and his social contacts. Also it must be admitted that Professor Peters has made no vital contribution to the difficult problem of organizing materials for a teaching course in educational sociology. Much of his material dealing with the family, the church, and other non-school educational agencies is either not germane or not so handled as to contribute to a science of education. Even those portions of the book, however, are worth-while reading for the prospective teacher and the teacher in service.

The author's great contribution lies in his insight into the scientific work which has been done in education and his insistence upon the research attitude in dealing with educational sociology. He has inserted innumerable question-problems into the body of the text and gives a list of suggestive research problems at the end of each chapter. The second part of the book is given over to the development of research techniques dealing with the determination of educational objectives and curricula. It is to be regretted that in this part, and in the book as a whole, so little attention is paid to the problems of admin-

istration and method, which are in as much need of sociological investigation as are those of objectives and the curriculum.

The book is well written, is filled with suggestive and significant materials, is thoroughly adapted to use as a textbook, and is certain to make a place for itself in the progressive schools of education.

WALTER R. SMITH

University of Kansas

BAGLEY, WILLIAM C., AND KEITH, JOHN A. H. *An Introduction to Teaching.* New York, Macmillan Company, 1924. 384 pp.

This volume is "the first of a series of books designed expressly for students of collegiate grade in professional schools for teachers." For high scholarship and professional sincerity it bears out the reputation of its authors. Other books in the list are *Junior High-School Curricula* by H. C. Hines, and *Procedures in High-School Teaching* by Waples. Still other volumes in this series are in preparation.

Many books have been written on the profession of teaching and unnumbered others will come from the presses to enlighten and inspire those who would do a great work well. There must be repetition, of course, in this multitude of books, but every capable author succeeds in saying something new or in interpreting in a new way the needs of the times. Bagley and Keith have succeeded in injecting into the text of their discussion a freshness and a timeliness which make their work eminently worth while. One can hardly conceive of their writing a book that would be dull, listless, and unprofitable.

The chapter titles are as follows: "Teaching as an Occupation;" "What is Teaching?" "The Materials of Education, The Basic Social Arts;" "The Materials of Education, The Liberal Studies;" "The Materials of Education, The Practical Arts and the Applied Sciences;" "Mind and Its Development;" "Nature's Provisions for Learning;" "Adapting Materials to Individual Capacities;" "The Outcomes of Teaching and Learning;" "Educational Systems;" "Personal Qualifications for Teaching;" and "Specific Qualifications for Specialized Fields of the Teaching Service." All the chapters are worth while and all make a distinct contribution to current educational thought.

An interesting feature of the volume is the set of questions, exercises, and problems at the end of each chapter. The lists of suggested references are up to date and cover the field of related educational thought excellently.

The book ought to find a hearty welcome in any school where teachers are trained, and it can render a distinct service in connection with the problem of training teachers in service.

LINNAEUS NEAL HINES

Indiana State Normal School, Terre Haute and Muncie

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

The Department of Attendance and Research of the Warren (Ohio) city schools has made an analysis of the costs of the schools of that city for the year 1923-24. In comparison with 73 cities of a population of thirty to one-hundred thousand, Warren spent slightly more for capital outlay, slightly less for debt service, and slightly less for current expenditures. Of the current expenses 75.6 percent were for instruction; 12.4, for operation; 4.2, for fixed charges; 3.6, general control; 2.3, for maintenances; and 1.9, for auxiliary agencies. Expenses are also analyzed for subjects and grades.

The Boston Department of Educational Investigation and Measurement repeated in October, 1924, the Courtis Tests which had not been given during the preceding five years. A comparison of the medians for 1924 and 1919, and the standards which were established on the basis of testing in 1915, was made. A comparison of these medians and standards show that in about one-half of the possibilities, the medians in 1924 are slightly higher than the standards, though slightly lower than the 1919 medians. In general, the situation has been maintained a little better in grades IV, V, and VI than in grades VII and VIII.

A study made by a committee on school problems of the Brooklyn (N. Y.) Teachers Association reveals a tremendous increase in the number of teachers per thousand who were absent on account of illness. In 1912-13, about 40 percent of the total number of teachers of the city were absent for one or more days; while in 1921-22, 78.7 percent were on the absent list. Classification by sex shows that the absence among women is more than twice as great as among men. Twenty-eight percent of the causes given was infectious diseases; 28 percent, diseases of the digestive system; 98 percent, diseases of the respiratory system; and 7 percent, diseases of the nervous system. The average duration of absence was 8.1 days for women, and 5.2 days for men. February is shown to be the month of greatest absence.

The *Pittsburgh Arithmetic Scales* by J. Freeman Guy has recently been distributed. The monograph presents a study in which previous arithmetic

scales are criticized and the following ten criteria set up for present work: equal difficulty, testing all four fundamentals in arithmetic at one period, reaching from low-third to high-eighth grade, composed of examples measuring from very little to very much achievement, having an easy yet adequate time control, easy to administer, readily scored, having marked diagnostic value, with facility of tabulation, produced at low cost. The study is Mr. Guy's Doctor's dissertation at the University of Pittsburgh, and is a valuable contribution to the field of arithmetic testing.

Superintendent J. C. Stiers of Washington, Pennsylvania, used, in October, a somewhat unique method to bring to the attention of his teachers one of the responsibilities of the administrative offices of the schools, and the modern theory of the function of education. He set up the following problem:

Fred J. left school at the end of the ninth grade. He worked in a carpenter shop for five months, clerked in a bookstore for six months, and then entered a newspaper office where he worked for thirteen months. The principal of the high school, Mr. Z., proposed that he return to school and graduate. Principal Z. said, "You have already earned four units, and I will allow you two units in manual training, two units in commercial work, one unit in salesmanship, one unit in printing, and one unit in business English. You may take five studies, English 11 and 12, Geometry, Physics, and American History and finish in a year." He graduated, entered college with a condition of two units of foreign languages, and finished with fair grades. He then entered business. Do you approve of the policy of the principal?

His teachers were divided about fifty-fifty on the answers, and Superintendent Stiers takes occasion, in a mimeographed bulletin, to analyze the situation and the schools' relation thereto.

The American Child Health Association is offering one thousand dollars in prizes to be divided among three schools making the most valuable contribution to a general study of successful health programs being carried on by secondary schools. The Association wishes as large a number of secondary schools as possible to report their work, giving, in as much detail as possible, the administration and supervision of the program, the course of instruction, how the work is correlated with other subjects, the connection with physical education, etc. Schools desiring to participate in the contest should file their reports before February 20, 1925. The schools east of the Mississippi River should place their applications with the Secretary of the Secondary School Health Program Study, American Child Health Association, 370 Seventh Ave., New York City; if west of the Mississippi River, the applications should be sent to American Child Health Association, 221 Sharon Building, 55 New Montgomery Street, San Francisco, California.

Mr. A. G. Yawberg, superintendent of Cuyahoga County (Ohio) schools and his assistant, Mr. Joseph A. Baer, have prepared an extensive bulletin on a plan for economy in school-building finance. It has a very attractive title, "Save Two Million Dollars on Better Schools in Cuyahoga County." The

bulletin takes into consideration the problems which will be faced by the county district during the next sixteen years, and presents a comprehensive program for meeting the needs of the schools during this time. It shows how the combination of districts will save large sums of money and yet result in increased educational opportunity. The proposal is also made for a thirty-dollar per-pupil state support from indirect taxes, a county district levy which when added to this thirty dollars will care for tuition and contingent, leaving each of the local districts with the necessity of raising all the money for buildings and sinking fund.

A bibliography of educational and psychological tests and measurements (Bulletin, 1923, No. 55, Government Printing Office in November, 1924), has been compiled by Margaret Doherty and Josephine MacLatchy, under the direction of Dr. B. R. Buckingham, director of the Bureau of Educational Research, Ohio State University. The bibliography is the most complete thus far published, and contains references to printed material issued from January 1, 1918, to June 30, 1922, so far as was obtainable. During the progress of the preparation of the bulletin a number of tests appeared, and some of the more important of these are also included. The material is organized under general references, references covering particular tests, references concerning the characteristics and uses of tests, uses according to types of schools, and a list of tests classified by subject.

The Association for Peace Education has published a bulletin on *An Analysis of the Emphasis upon War in our Elementary-School Histories.* The study was made by Professor Monroe of Carleton College, Professor McElhammon of Baylor College, and Ralph Henry of Carleton. A quantitative study was made of twenty-four textbooks from the standpoint of word count and illustrative content. The word count was found to vary from 19 to 43 percent pertaining to war causes, activities, and results; while the content pertaining to peace varied from zero to 16 percent. The illustrations for war varied from 14 to 67 percent and for peace from zero to 8 percent. Stevenson's *A School History of the United States,* published in 1921, according to this analysis emphasizes peace in both word count and content more than any other text. Even so, the proportion devoted to war is more than twice as much in word count and nearly four times as much in illustration as the proportion devoted to peace. An analysis of supplementary readers, by the same criteria, reveals about the same relative emphasis, as found in the history textbooks. Qualitative study of the same books likewise showed great variations.

Miss Ruth Gillette Hardy, a teacher at the Girls' Commercial High School, New York City, after a year of travel and observation of schools in other parts of the world, has written a brief article for the *Schoolmaster,* the official organ of the National Union of Teachers of the British Isles. Among the contrasts pointed out are: (1) British schools are smaller than American schools, for no school she visited had more than six hundred pupils. (2) All

head masters are required to do some teaching. (3) The opportunities for secondary education for poor boys are much fewer than in America. (4) The twin curses that economy has laid upon British schools are oversize classes and limited educational opportunity. One paragraph of this article reads:

If I balance our alternating strength and weakness in English schools as compared with those of the United States, I should sum up that the schools of England, great and small, sacrifice the average young person to the able, so that in the fifty years that you have had public education you have developed national leaders from all walks of life but they have difficulty in keeping in touch with the ordinary citizen; the schools of the United States have sacrificed the able to the average, so that we have a citizenry of high average education, but tremendous difficulty in finding leaders. When we both have proper school systems we shall find some way not to sacrifice either.

Dr. James E. Lough of New York University urges the annual rating of teachers in any school system on the basis of the following five principles:

1. To disclose to the teacher her high points and her low points for the purpose of enabling her to overcome such weaknesses as interfere most seriously with her success, and to employ her strong points to the best advantage.
2. To enable a teacher to know definitely how her work stands in relation to the work of other teachers in the same school system.
3. To keep clearly before the mind of the teacher a definite conception of the professional qualities she must possess.
4. To enable superintendents and supervisors to judge the work of teachers impersonally, applying the same standards of measurement to all alike.
5. To enable administrative officers to differentiate teachers according to general or special aptitudes and so to assign properly qualified teachers to positions developing special traits.

The Question of Class-Size

The problem of what is the optimum size of class is one of great importance in the organization of school systems. During the last fifteen years a number of more or less scientific investigations of this question have been made. The general plan of procedure, especially in the more recent investigations, has been to endeavor to hold constant all factors, except that of size of class, and to base the conclusions reached upon the measured achievements of the pupils in classes of different sizes. Most of the evidence secured has indicated that size of class is not so important a factor in the achievements of the pupils as many have supposed. In connection with a study made chiefly for another purpose, the writer collected data which tend to support the conclusion that class-size is relatively unimportant. According to them, the relationship existing between class-size and certain other measures of school efficiency appears to be very small. The study referred to included some 53,000 pupils in the elementary schools of eighty-two cities and towns in Illinois. The average size of elementary class was computed for each school system. Among the other items found were the average progress, that is, the average rate at which pupils had progressed through the grades since entering the system, the average retardation, computed in the usual manner, and indices

of holding power which showed the percents that the twelve- and thirteen- and the fifteen- and sixteen-year-old pupils were of the standard age group. It was found that the coefficients of correlation between the average class-size and each of the other measures just mentioned were 0.20 or less. As the probable errors of these coefficients were each .07 the conclusion to be drawn is that there was very little if any connection between the average size of class and average progress, average retardation, or holding power.

C. W. ODELL

University of Illinois

Harold—a Second Report

A brief account of Harold M. Finley has already appeared.[1] Tested with the Stanford-Binet at the age of four years and two months (April 15, 1920) he earned an I.Q. of 168. At that time he read fluently in a number of first-grade readers at sight and in many other ways demonstrated his undoubted intellectual superiority.

Accounts of Harold's progress which reached us from time to time indicated rapid advancement and a sustained interest in intellectual pursuits. In September, 1922, at the age of six years and seven months, Harold presented himself for admission at the local schools. The superintendent had anticipated his coming and during the summer had been much concerned as to what he should do with him. Advice was sought and easily secured, but in the minds of those discussing the case there was marked disagreement as to what constituted sound educational procedure in dealing with children as bright as Harold. The superintendent was much perturbed by the "accelerated-progress-versus-enriched-curriculum" controversy.

It was tentatively decided to start Harold in the fourth grade. It soon became apparent to all concerned, however, that it would be impossible to keep him there for the work which his teacher was able to provide for him proved to be very easy and occupied but a small portion of his time. Accordingly, he was advanced to the fifth grade at the end of the first month where he spent the remainder of the year. But fifth-grade work did not prove to be a fair challenge for his ability so he was permitted to skip the sixth grade and in September, 1923, began his second year in school by entering grade seven. Harold spent the year in grade seven taking considerable pride in making a perfect record for punctuality and attendance.

During the year nearly all of his grades were E's. He had one G in history and a number of G's in language. In the latter subject the pupils in this system are required to diagram sentences and parse words. Having skipped the sixth grade, Harold was handicapped in this work but eventually mastered it and made 100 percent on his final examination. He was the only one in his class to make 100 percent on the final examination in arithmetic.

These facts we had in our possession when Harold, accompanied by his parents, visited us on June 6 and 7, 1924. We shall report here the results of the tests which we were able to give him at that time.

[1] "A Bright Boy," *Journal of Educational Research*, 2:865-68, December, 1920.

First, we repeated the Stanford-Binet, starting with Year X. To save time, only the starred exercises in Terman's record booklet were used. He passed successfully all of the tests given him for Year X and Year XII, all but the vocabulary of Year XIV and Year XVI, and passed one, repeating seven digits backwards, in Year XVIII. This gave him a mental age of sixteen years and three and one-half months and an I.Q. of 197. In the arithmetical reasoning test of Year XIV where a maximum time of one minute is allowed for each of the three problems, two of which must be correct if the test is passed, Harold gave the answers to the three problems correctly using about one second on each.

After a rest period Harold took the Stanford Achievement Test, Advanced Examination, Form A. This test is designed for grades IV-VIII. In the following table we give his scores on the various subtests and, as far as possible, the grade and educational age equivalents as supplied by the revised table of norms issued in July, 1924. This table of norms gives grade equivalents up to 10.0 and educational ages up to 17.7, above which they are designated as "adult."

TABLE I.—HAROLD'S SCORES ON THE STANFORD ACHIEVEMENT TEST

Test	Score	Grade	Educational Age	Norms for Grade X
1. Reading: Paragraph Meaning	88	Above 10.0	Adult	86
2. Reading: Sentence Meaning	52	8.0	13.8	64
3. Reading: Word Meaning	70	Above 10.0	Adult	67
4. Arithmetic: Computation	168	Above 10.0	Adult	142
5. Arithmetic: Reasoning	128	Above 10.0	Adult	104
6. Nature Study and Science	70	9.5	17.2	74
7. History and Literature	47	8.0	13.8	66
8. Language Usage	34	8.5	14.5	41
9. Dictation Exercise	186	Above 10.0	Adult	176
Composite Score	84	Above 10.0	Adult	82

In the test of reading for paragraph meaning, it will be noted that Harold made a score of 88 which places him above the tenth-grade level and gives him an adult educational age. The table shows that in five of the nine subtests Harold scored above the tenth-grade norm and that his lowest scores are just equivalent to the norms for pupils entering grade eight.

To help the reader in interpreting Harold's performances in the five subtests in which he scored higher than the tenth-grade norms, we give the norms for grade 10.0 in the last column of the table. His most outstanding performances, it will be noted, are in arithmetic. In computation he exceeds the tenth-grade norm by 18 percent of that norm and in reasoning by 23 percent.

The following day, Harold was given a miscellaneous assortment of tests in an effort to further explore his abilities and as a partial check upon the findings reported above.

On the Ohio Literacy Test, a five-minute test of fifty questions to be answered by "Yes" or "No," he scored 29. This is equivalent to the

eighth-grade norm and to a mental age of 13 years. Not a single error was made on this test. Harold worked rather deliberately making a perfect score so far as he went.

On the Porteus Maze Tests he obtained a test age of 13 years. The test for the ten-year level required two trials but all others up to the twelve-year level were successfully negotiated on the first trial. The twelve-year test required four trials, there being considerable evidence of impulsiveness. He succeeded in one trial on the fourteen-year test, the highest of the series.

On the Stenquist Assembly Test, Harold made a very poor showing. When confronted with this test he immediately stated that he did not think that he could do much with it for, he said, "I am not very good at putting things together." It was quite evident that his mechanical ability was of a far lower order than his other abilities which had been explored.

Partly to satisfy the examiner's curiosity, Harold was given the Ohio State University Psychological Test, Form 3, a general intelligence test which is given to all Freshmen entering Ohio State University. Out of a possible 230 points, he scored 77. As measured by norms from 2,200 Freshmen, this is an 18-percentile score. In other words, his score exceeded that of 18 percent of the Ohio State University Freshmen.

A sample of Harold's handwriting, produced with a pen, was scored by three judges independently. The mean of the three ratings on the Ayres Scale was 68, six points higher than the eighth-grade norm as given by Ayres. The speed of his writing was not measured.

The Porteus and Stenquist Tests were administered by Mr. L. N. Yepsen, now of the Training School of Vineland, New Jersey. The following anthropometric data were also collected by Mr. Yepsen.[2]

Weight	24.4 kg	or 51-percentile
Standing height	1290 mm	or 83-percentile
Sitting height	700 mm	or 80-percentile
Right grip	15 kg	or 70-percentile
Left grip	13 kg	or 50-percentile
Lung capacity	1300 cc	or 40-percentile
Head length	180 mm	
Head breadth	140 mm	
Head height	131 mm	
Calculated cubic capacity of head	1324 cc	or 68-percentile

Harold has recently acquired a typewriter and has taken it up seriously. A court stenographer of the family's acquaintance gave him a lesson or two and supplied him with a chart showing the proper fingering. He already shows fair proficiency.

At the present time Harold shows very much interest in writing. He published a story in *Child Life*, December, 1923, page 875. He has recently written two stories of much greater length which he is very desirous of publishing in booklet form. They contain some grammatical inaccuracies and rhetorical inelegancies, to be sure, but suggest valuable imaginative qualities.

[2]There is a picture of Harold in the *American Magazine* for April, 1922, p. 62.

Harold's interests are, of course, on a plane much higher than those of most eight-year-olds. He reads a great deal, particularly books having to do with nature. He plays too; he says that he likes to climb, but the usual games of children of his age do not interest him. In general, he does not care for movies but the antics of Jackie Coogan always entertain him.

Harold is much interested in the affairs of the church to which the members of the family belong. After much importuning he was permitted to lead the mid-week prayer meeting on June 3, 1924. He prepared his plans with great care, persuading a number of people to make short talks on topics which he assigned. He prepared and made his own talk and offered a prayer. The meeting lasted an hour and ten minutes and at no time were matters allowed to drag. Needless to say, the church was crowded.

Although the intelligence quotient which we obtained with the Stanford-Binet may be too high since abbreviating this test probably reduces its reliability, yet after examining the data which we were able to collect we are persuaded that the intelligence quotient which we originally secured, 168, is too low. We shall follow Harold's development with much interest and shall subsequently give further reports as to his progress.

R. L. MORTON

Ohio State University

The College of Education, Ohio State University, will hold its Fifth Annual Educational Conference at Columbus, Ohio, on April 2, 3, and 4. As in past years, the work of the Conference will center on a general theme. The keynote for this year has been announced as "Democracy and Education." Three general programs and twenty-six sectional meetings, all focusing on the Conference keynote, will be held during the three-day gathering.

Speakers of national reputation are being negotiated with for both the general and sectional meetings. In all, more than one hundred papers will be presented at the Conference, and in the sectional meetings, round-table discussions of the speeches will be provided for. The sections will meet on the second day of the Conference, Friday, April 3.

The Ohio Conference is gradually becoming something more than a state-wide institution. In 1924, delegates from Michigan, Indiana, and West Virginia attended the Fourth Annual Conference. Reports reaching the Executive Committee for the 1925 Conference indicate that these states, and others, will be represented at Columbus this year.

Members of the Executive Committee for the Fifth Annual Ohio State University Educational Conference are: George F. Arps, dean of the College of Education, *chairman;* P. R. Stevenson, *secretary;* B. R. Buckingham, J. C. Morrison, O. G. Brim, B. H. Bode, E. A. Doll, and H. G. Hullfish.